

**TOWARDS CONDITIONAL GENE  
INACTIVATION IN THE MOUSE COLON  
AND  
FUNCTIONAL ANALYSIS OF GENES  
INFLUENCING COLORECTAL  
CARCINOGENESIS**

**Jason Tyler Read**

**Submitted for Degree of Doctor of Philosophy**

**University of Edinburgh**

**2000**



## **Declaration**

This thesis and the work described herein is solely my own work. Where contributions have been made by others this is explicitly stated in the text.



## Contents

Title page	1
Declaration	2
Contents	3
Acknowledgements	7
Abbreviation	8
Amino Acid Abbreviations	10
Abstract	11
Chapter 1 - Introduction	13
1.1 Human Colorectal Cancer	14
1.1.1 Incidence and Epidemiology	14
1.1.2 Molecular Biology and Genetics of Colorectal Cancer	15
1.1.2.1 Hereditary Non-Polyposis Colorectal Cancer (HNPCC)	15
1.1.2.2 Familial Adenomatous Polyposis (FAP)	16
1.2 The Adenomatous Polyposis Coli ( <i>APC</i> ) Gene	17
1.2.1 Genomic Organisation in Mouse and Human	17
1.2.2 Mutations of the <i>APC</i> gene	20
1.2.3 The APC tumour suppressor protein; Structures and Protein Interactions	20
1.2.4 Expression pattern	27
1.2.5 Function	28
1.3 The Wnt/Wingless signalling pathway	29
1.3.1 Components of the Wnt Pathway	29
1.3.1.1 Wnt	29
1.3.1.2 Frizzled	29
1.3.1.3 Dishevelled	30
1.3.1.4 Glycogen Synthase Kinase 3 $\beta$ (GSK-3 $\beta$ )	31
1.3.1.5 Axin	32
1.3.1.6 $\beta$ -catenin	35
1.3.1.6.1 Structure and function	35
1.3.1.6.2 Downstream targets of $\beta$ -catenin	37
1.3.2 The Wnt pathway in development	39
1.3.3 Dysregulation of the Wnt /Wingless pathway	43
1.4 The <i>Min</i> Mouse; a murine model of FAP	45
1.4.1 Phenotype of the <i>Min</i> Mouse	45
1.4.2 Embryonic lethality of the homozygote	47
1.4.3 Comparison of the <i>Min</i> model and human disease	47
1.5 Transgenic and knock-out mice as cancer models	48
1.5.1 Introduction of exogenous DNA into mouse oocytes; transgenics	50
1.5.2 Site specific alteration of the mouse genome to produce disease models; Gene targeting	51
1.5.3 Conditional Gene Targeting	55
1.5.3.1 Background: The FLP/frt system	55
1.5.3.2 Background: The Cre/LoxP system	56
1.5.3.3 Manipulation of the genome using the Cre/LoxP system	57

1.5.3.4 Regulating the expression and/or activity of the Cre recombinase	63
1.5.3.5 Site-specific insertion	66
1.5.3.6 Site specific excision of LoxP flanked DNA sequences	67
1.5.3.6.1 Activation of expression through excision (Floxed stop)	67
1.5.3.6.2 Inhibition of expression through excision (Gene deletion)	68
1.6 Project Aims	69
<b>Chapter 2 - Materials and Methods</b>	<b>71</b>
2.1 Basic cloning Methodology	72
2.1.1 Agarose Gel Electrophoresis of DNA	72
2.1.2 Restriction Digest Analysis of Plasmid DNA	72
2.1.3 Extraction of DNA Fragments from Agarose Gels	73
2.1.4 Dephosphorylation of DNA fragments	74
2.1.5 Blunting of restriction generated overhangs	74
2.1.6 DNA Ligation	75
2.1.7 Transformation of Bacteria with Plasmid DNA	76
2.1.8 Preparation of Plasmid DNA	77
2.1.8.1 Small Scale Plasmid Preparation	77
2.1.8.2 Large Scale Plasmid Preparation	78
2.1.8.3 Preparation of BAC DNA	79
2.1.8.4 Preparation of bacterial glycerol stocks	80
2.1.9 Sequencing of Plasmid/PCR DNA	80
2.1.10 Southern Blot/Dot Blot hybridisation	81
2.2 PCR reactions	83
2.2.1 Reaction Conditions	83
2.2.2 Primer Design	84
2.3 Tissue culture	87
2.3.1 Derivation of cell lines	87
2.3.2 Thawing, media and incubation conditions	88
2.3.3 Harvesting cells for protein	90
2.3.4 Chemical treatment of cells	90
2.4 Immunoblotting	91
2.4.1 Tissue sample preparation	91
2.4.2 Protein quantification	91
2.4.3 Protein electrophoresis	92
2.4.4 Protein Transfer	92
2.4.5 Antibody incubation	93
2.4.6 Signal detection	93
2.7 Introduction of DNA into cultured mammalian cells	94
2.7.1 Microinjection	94
2.7.1.1 DNA preparations	94
2.7.1.2 Injection conditions	94
2.7.2 Transfection	94
2.7.2.1 Calcium phosphate transfection	94
2.7.2.2 Liposome mediated transfection	95

2.8 Animal work	96
2.8.1 Tissue processing	96
2.8.2 Timed matings	97
<b>Chapter 3 - Time and tissue-specific gene deletion</b>	<b>98</b>
3.1 Background	99
3.2 Isolation of <i>Cdx-1</i> tissue-specific promoter sequence	103
3.2.1 BAC screening	103
3.3 Construction of BAC based targeting and transgenic vectors	104
3.3.1 Subcloning of CreER <sup>tm</sup> coding sequence	111
3.3.2 PCR generation of mutant floxed and native kanamycin/neomycin resistance cassettes	114
3.3.2.1 <i>in vitro</i> Cre recombination	120
3.3.3 PCR and cloning of arms of homology	122
3.3.4 Construction and screening of final BAC targeting vectors	126
3.4 Discussion	140
<b>Chapter 4 - Evaluation of murine <i>Cdx-1</i> promoter sequence</b>	<b>142</b>
4.1 Background	143
4.2 Mapping of newly isolated upstream <i>Cdx-1</i> promoter sequence	147
4.3 Subcloning and sequencing of pGEM-XhoI	148
4.4 Reporter plasmid construction	152
4.4.1 Construction of pEGFP-Sph	153
4.4.2 Construction of pEGFP-LongEco and pEGFP-ShortEco	157
4.4.3 Construction of pEGFP-SCdx	157
4.5 Introduction of pEGFP- <i>Cdx-1</i> plasmids into mammalian cells	164
4.5.1 Microinjections	164
4.5.2 Transfections	168
4.5.2.1 HCT 116 cell line	169
4.5.2.2 SW 480 cell line	172
4.5.2.3 Primary murine colonocytes	174
4.5.2.4 Murine HM-1 ES cells	177
4.6 Discussion	185
<b>Chapter 5 - Characterisation of a transgenic Floxed <i>APC</i> mouse line</b>	<b>193</b>
5.1 The floxed <i>APC</i> line	194
5.2 Derivation of floxed <i>APC</i> line	195
5.2.1 <i>in vitro</i> analysis of transgene expression	198
5.3 Analysis of the <i>CC</i> transgenic mouse line	200
5.3.1 <i>Min</i> rescue	205
5.3.2 Comparison of survival times	206
5.3.3 Comparison of tumour burden	211
5.3.4 Comparison of histological sections	214
5.3.5 Transgenic embryos	214
5.4 Discussion	218

<b>Chapter 6 – Methylation and manipulation of the Wnt-1 pathway</b>	<b>221</b>
6.1 Background	222
6.2 <i>in vivo</i> MNNG treatment	224
6.3 Embryonic fibroblasts	226
6.3.1 Preparation and genotyping	226
6.3.2 MNNG induced alteration of $\beta$ -catenin levels	226
6.3.2.1 $\beta$ -catenin levels in wild type fibroblasts	226
6.3.2.2 $\beta$ -catenin levels in p53 <sup>-/-</sup> fibroblasts	227
6.3.2.3 $\beta$ -catenin levels in <i>Msh2</i> <sup>-/-</sup> fibroblasts	227
6.3.2.4 $\beta$ -catenin levels in p53 <sup>-/-</sup> / <i>Msh2</i> <sup>-/-</sup> fibroblasts	227
6.4 Discussion	232
<b>Chapter 7 – Conclusions</b>	<b>235</b>
<b>References</b>	<b>241</b>
<b>Appendix 1 – Solutions</b>	<b>267</b>
<b>Appendix 2 – <i>Cdx-1</i> promoter sequence data</b>	<b>276</b>
<b>Appendix 3 – CC mouse database</b>	<b>285</b>

## **Acknowledgments**

Thanks first of all to my supervisors, Prof. Martin Hooper and Dr. Jane Armstrong, for ideas, guidance and above all patience. I realise I'm not the easiest writer to proofread...I blame the french education.

Thanks to Tim for the flow skills (but especially for the sympathetic ear) and to Owen for the constant supply of fibroblasts and insight.

Special thanks, of course, to the United Transgenics Collective and the residents of A624 (past and present) for the intensive discussions regarding tiny dolphins, the feasibility of lab work on a Cuban beach, glowing pets of all manner, the effects of lunchtime pints, the lizard as a model organism for the study of oxidative stress, the existence of beards, extravagant lunches, "funny results" which were patently not funny and the indepth (and ongoing) study into the effects of caffeine on the human nervous system. While these subjects have not been covered in my thesis they were vital to giving me the mental strength to finish it.

Finally thanks to Anita for keeping me together, even from the other side of the Atlantic, and to Mom and Dad for pretty much everything.

## Abbreviations

APC	adenomatous polyposis coli
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BAC	Bacterial Artificial Chromosome
bp	basepair
BRL	Buffalo Rat liver
BS	Brain Specific
BSA	Bovine serum albumin
CBP	CREB-binding protein
CKII	casein kinase II
CREB	cAMP responsive element
CTP	Cytosine triphosphate
ddATP	dideoxyadenosine triphosphate
ddCTP	dideoxycytosine triphosphate
ddGTP	dideoxyguanosine triphosphate
ddNTP	dideoxynucleoside triphosphate
ddTTP	dideoxythymidine triphosphate
dH <sub>2</sub> O	deionised water
DMSO	Dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNA-PK	DNA-dependant protein kinase
dNTP	deoxynucleoside triphosphate
DTT	Dithiothreitol
EDTA	disodium ethylenediamine tetraacetate
EGFP	Enhanced Green Fluorescent Protein
ENU	EthylNitrosourea
ER	Oestrogen Receptor
ES cell	Embryonal Stem Cell
EtBr	Ethidium Bromide
FAP	Familial Adenomatous Polyposis
FCS	foetal calf serum
flox	flank by loxP
G418	geneticin sulphate
GFP	Green Fluorescent Protein
GSK-3 $\beta$	glycogen synthase kinase 3 $\beta$
GTP	guanosine triphosphate
H&E	Haematoxylin and Eosin
HBSS	Hanks' Balanced Salt Solution
HEPES	N-2-Hydroxyethyl piperazine-N'-2-ethanesulphonic acid
HMG	High Mobility Group

HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HR	homologous recombination
HSP	heat shock protein
HSV	Herpes Simplex Virus
<i>Kan</i>	Kanamycin
kb	kilobasepair
kDa	kiloDalton
LB	Luria Broth
LBD	Ligand Binding Domains
LEF	Lymphocyte enhancer factor
LIF	Leukaemia Inhibitory Factor
LoxP	Locus of crossover of P1
MEF	Murine Embryonic Fibroblast
<i>Min</i>	multiple intestinal neoplasia
MMR	Mismatch Repair
MMTV	Mouse Mammary Tumour Virus
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
MOPS	3- [N-morpholino]propanesulphonic acid
<i>MSH2</i>	MutS homologue 2
NCS	newborn calf serum
<i>Neo</i>	Neomycin/Kanamycin resistance gene
PARP	poly-ADP ribose polymerase
PCR	Polymerase Chain Reaction
PGK	3-phosphoglycerate kinase-1
Pla2g2	secretory phospholipase-a2
PP2A	protein phosphatase 2a
RA	Retinoic Acid
RARE	retinoic acid response element
RNA	ribonucleic acid
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SDS	Sodium Deodecyl sulfate
SMAD	Small/Mothers Against Dpp
SSC	standard saline-citrate
STAT	Signal Transducer and Activator of Transcription
SV40	Simian Virus 40
TBE	Tris Borate EDTA (see appendix 1)
TCF	T-cell factor
TE	Tris EDTA solution (see appendix 1)
TEMED	N,N,N,N'-tetramethylethylenediamine
Tk	thymidine kinase
TTP	thymidine triphosphate
X-gal	5-Bromo-4-Chloro-3-Indolyl- $\beta$ -D-Galactopyranoside

## List of Amino Acid abbreviations

Alanine	A or Ala
Arginine	R or Arg
Asparagine	N or Asn
Aspartic Acid	D or Asp
Cysteine	C or Cys
Glutamine	Q or Gln
Glutamic Acid	E or Glu
Glycine	G or Gly
Histidine	H or His
Isoleucine	I or Ile
Leucine	L or Leu
Lysine	K or Lys
Methionine	M or Met
Phenylalanine	F or Phe
Proline	P or Pro
Serine	S or Ser
Threonine	T or Thr
Tryptophan	W or Trp
Tyrosine	Y or Tyr
Valine	V or Val



## Abstract

Dysregulation of the Wnt signalling pathway is implicated in the carcinogenic transformation of normal cells in a number of human tissues. Mutations in Wnt pathway signalling molecules appear to be particularly important in the development of colorectal neoplasia. This thesis investigates the effects of modulating levels of Wnt pathway components on downstream molecules *in vivo* and *in vitro*.

To examine the effects of loss of Wnt pathway components on the biology of the mouse large intestine work was undertaken to assemble a tissue-specific, inducible Cre (CreER<sup>tm</sup>) transgene construct. This formed part of a group effort using various lengths of upstream sequence from the mouse homeobox gene *Cdx-1* to drive CreER<sup>tm</sup>. In order to ensure that sufficient *Cdx-1* promoter sequence to achieve adequate expression levels and localization, a recently described methodology based on the Red $\alpha$ - $\beta$ - $\gamma$  recombination system of bacteriophage Lambda was used in an effort to insert the CreER<sup>tm</sup> coding sequence into a Bacterial Artificial Chromosome (BAC).

The BAC used in the construction of the transgene was found to contain previously uncharacterised upstream *Cdx-1* promoter elements, which may play a role in determining correct tissue expression pattern. In order to further characterise the newly isolated *Cdx-1* promoter elements, portions of the BAC derived sequence were used in the construction of reporter plasmids encoding the Enhanced Green Fluorescent Protein (EGFP). The EGFP based reporter plasmids, which contained various fragments of the *Cdx-1* promoter encoding a number of transcriptional control elements, were microinjected and liposome transfected into murine embryonic stem (ES) cells, colorectal cell lines and primary murine colonocytes in order to determine the contribution of the transcriptional elements to levels of expression.

One of the primary targets for gene excision using a large intestine specific, inducible Cre recombinase would be the Adenomatous Polyposis Coli (*APC*) gene, the *APC* gene product being a central mediator of the Wnt pathway.

Mutations of the *APC* gene occur in a large majority of spontaneously arising intestinal tumours in human patients and germline *APC* mutations are responsible for the human condition Familial Adenomatous Polyposis (FAP). A previously generated floxed *APC* transgenic mouse strain was crossed onto a mouse line bearing a germline truncation mutation (*Min*) in *APC*. The resulting progeny were characterised to determine if the introduced transgenic material had an influence on disease phenotype; comparison of survival times and embryogenesis indicate that the transgenic material has negligible effect on the *Min* phenotype.

Another factor contributing to the development of colorectal tumours in humans appears to be deficiencies in DNA mismatch repair processes. The relationship between levels of the Wnt pathway molecule  $\beta$ -catenin following methylation damage and the functionality of the DNA damage sensor p53 and DNA mismatch sensor *Msh2* were explored *in vitro* in Murine Embryonic Fibroblasts (MEFs).  $\beta$ -catenin levels were found to drop following application of the methylating agent MNNG in wild type MEFs but remained unchanged in cells containing mutant p53. Cells null for *Msh2* function showed the same drop in  $\beta$ -catenin levels as wild type fibroblast, suggesting that *Msh2* is not involved in Wnt pathway reaction to methylation damage.

The work presented in this thesis contributes to the development of a transgenic mouse line bearing a conditional, inducible intestinal epithelial specific Cre recombinase and to the understanding of regulation of the Wnt signalling pathway in response to genotoxic damage as well as expression of the murine homeobox gene *Cdx-1*.

# **Chapter 1**

## **Introduction**

## **Chapter 1 – Introduction**

### **1.1 Human Colorectal Cancer**

#### **1.1.1 Incidence and Epidemiology**

One of the most common tissues to become cancerous in the human is the large intestinal epithelium. A high rate of cellular turn over requiring rapid cell division coupled with the presence of potentially damaging chemical factors in ingested substances combine to give a relatively elevated rate of neoplasia. While cancer of the small intestine is rare in humans, in the developed nations of the world the incidence of colorectal cancers is such that 50% of individuals can expect to have developed a tumour by the age of 70. In the same geo-political areas colorectal cancer is the second most common cause of cancer deaths and is in fact the most common form of non-smoking related cancer (Parker *et al.*, 1996). With populations ageing and more and more societies adopting a “western” style diet which predisposes to colorectal tumour formation, it is reasonable to believe that the number of people afflicted with this disease will continue to rise worldwide (Giovannucci and Willett, 1994; Burnstein, 1993). With these consideration it is not surprising that cancers of the large intestine, with special emphasis on cancers of the colon, have received a great deal of attention and been the focus of much intensive research.

Early on in the investigation of colorectal cancer it was noticed that a significant proportion of these cancers (estimates range from 5-15%) appeared to occur within families, susceptibility to the development of colorectal tumours being inherited in a dominant fashion (Cannon-Albright *et al.*, 1988; Houlston *et al.*, 1992a). Through study of the genes, proteins and pathways involved in the increased rates of colorectal cancers in these syndromes a greater understanding not only of this specific cancer but also of the process of cancer formation as a whole has been achieved.

### **1.1.2 Molecular Biology and Genetics of Colorectal Cancer**

As mentioned above there has been a great deal of research looking for common genetic defects which may predispose to development of intestinal cancers. To date two syndromes have been defined, both leading to greatly increased rates of tumour formation; the first being Hereditary Non-Polyposis Colorectal Cancer (HNPCC) and the second being Familial Adenomatous Polyposis (FAP). HNPCC patients account for up to 6% of the hereditary predispositions to colorectal cancers while FAP accounts for only 1% of these individuals (Salovaara *et al.*, 2000; Aaltonen and Peltomaki, 1994).

#### **1.1.2.1 Hereditary Non-Polyposis Colorectal Cancer (HNPCC)**

The investigation of the nature of the inherited mutation (or mutations) involved in HNPCC proved to be more difficult than expected. While it was easily determined that the syndrome was being inherited in a Mendelian fashion, the search for the common mutated allele proved difficult. The unravelling of the problem began when the search for lost microsatellite marker alleles, presumably deleted along with the apparent missing tumour suppressor gene turned up changes in the lengths of several microsatellites rather than deleted ones, a defect termed Microsatellite Instability (Aaltonen *et al.*, 1993). This suggested that the mutation was not simply deleting the functions of a traditional tumour suppressor gene but was instead giving rise to changes within the DNA sequence, which may have deleterious effects on actual tumour suppressors. Studies of DNA repair in bacteria had already revealed mutant cells with an increased rate of mutation; the mutation leading to this phenotype was in one of a group of genes involved in the repair of mismatched bases, the result of some sort of genotoxic insult (Strand *et al.*, 1993). The 5 human homologues of these genes, the Mismatch Repair (MMR) genes, proved to be the sites of the heritable mutations which lead to HNPCC (Fishel *et al.*, 1993; Strand *et al.*, 1993; Boland *et al.*, 1998). Today it is known that the majority of the mutations leading to HNPCC are in three MMR genes; hMSH2, hMLH1 and hPMS2. As mentioned previously the intestinal epithelium is a tissue in constant renewal, its cells dividing and being shed constantly. It is not surprising then that germline

mutations in MMR genes would show themselves first in this tissue; the combination of a potentially damaging chemical environment with the need to constantly replicate DNA would lead to a rapid accumulation of mutations throughout the genome, some, eventually, in tumour suppressors or proto-oncogenes.

### **1.1.2.2 Familial Adenomatous Polyposis (FAP)**

Unlike HNPCC, where the intestine remains free of polyps prior to the formation of disease the hereditary disorder familial adenomatous polyposis (FAP) is characterised by the development of hundreds to thousands of adenomas of the colorectal mucosa. Eventually one or more of these regions of abnormal cell growth and turnover progress to a cancerous stage (Cannon-Albright *et al.*, 1991; Cannon-Albright *et al.*, 1988). FAP patients may also display a congenital hyperplasia of the retina, indeed this may be the first apparent symptom of an FAP genotype (Houlston *et al.*, 1992b). Again in contrast to the case of HNPCC, the genetic defect responsible for this disorder was relatively quickly mapped and the mutations responsible identified. In the large majority of cases of FAP this debilitating phenotype has been traced to a germline mutation in the adenomatous polyposis coli (*APC*) gene (Nishisho *et al.*, 1991; Groden *et al.*, 1991; Kinzler *et al.*, 1991a; Kinzler *et al.*, 1991b). It was soon discovered that a large majority of sporadic colorectal tumours (adenomas and cancers), up to 80%, were found to be missing alleles of this gene and/or to contain mutated sequences coding for truncated versions of the APC protein (Ichii *et al.*, 1992). These mutations of the *APC* gene were soon implicated in a large number of tumours derived from other tissues, including osteomas, desmoid tumours and breast tumours (Burt *et al.*, 1992; Davies *et al.*, 1995; Miyaki *et al.*, 1993; Szabo and King, 1995).

The role of *APC* gene mutations in the formation of a number of tumour types prompted intensive research into the nature of the *APC* gene and its protein product. Consequently the role the *APC* gene product plays in the maintenance of regular cell activities has been further understood yet the implications and end effects of the pathways involving APC have yet to be fully clarified.

## **1.2 The Adenomatous Polyposis Coli (APC) Gene**

### **1.2.1 Genomic Organisation in Mouse and Human**

The *APC* genes of the human and the mouse occur in relatively conserved regions. In the human the *APC* gene is found on chromosome 5 at q12 while in the mouse the homologue is to be found on chromosome 18 also at q12 (Kinzler *et al.*, 1991a; Kinzler *et al.*, 1991b; Luongo *et al.*, 1993). Originally the gene was thought to be composed of 15 exons with the final exon comprising approximately half of the coding sequence (Santoro and Groden, 1997). A number of additional exons have since been discovered, several of which are differentially expressed and represent alternative splice forms (see figure 1.1) (Santoro and Groden, 1997; Samowitz *et al.*, 1995; Thliveris *et al.*, 1994). The best studied alternative APC isoform is one containing the Brain Specific (BS) exon, now known not to be restricted to expression in the brain, found upstream of the originally identified exon 1; this exon is translated to the exclusion of exon 1 in differentiated tissues such as the brain (Pyles *et al.*, 1998). At least three exons in addition to the BS exon have been found upstream of the originally identified exon 1, these have been termed (from 5' to 3') exons 0.3, 0.1 and 0.2; transcripts containing these alternative upstream exons have been identified both with and without exon 1 (Horii *et al.*, 1993; Thliveris *et al.*, 1994). Interestingly exon 1 contains a translation stop codon upstream of its translation start codon suggesting that transcripts beginning with an alternative upstream codon which contain exon 1 produce truncated proteins and that only those transcripts containing an alternate upstream exon which lack exon 1 can produce full length proteins (Thliveris *et al.*, 1994; Santoro and Groden, 1997; Pyles *et al.*, 1998). Additionally exons 9 and 10 have alternative splice forms (9A and 10A) bringing the total number of exons to 21 giving rise to at least 16 different transcripts, including some lacking exon 7 (Santoro and Groden, 1997; Oshima *et al.*, 1993).

**Figure 1.1:** A schematic representation (not to scale) of the exonic/intronic organisation of *APC*. The figure shows the large number of alternatively expressed exons as well as representing the various possible alternative splicing forms.





### **1.2.2 Mutations of the APC gene**

As mentioned previously, many intestinal tumours arising spontaneously, in addition to those tumours developed by FAP individuals have been found to contain mutations of the *APC* gene. Screening of the *APC* mutations from these tumours has produced a map of mutation frequencies for the length of the *APC* gene (Krawczak and Cooper, 1997). The large majority of disease causing mutations have been found to be those resulting in truncated reading frames, including stop mutations introduced through transposition mutations as well as frameshift mutations giving rise to premature stop codons (Mandl *et al.*, 1994). While mutations have been found throughout the length of the *APC* coding sequence the majority of mutations resulting in deleterious phenotypes, in both germline and somatic scenarios, occur in the 5 prime portion of the coding sequence. Additionally in germline mutations there appear to be two sites of increased mutation frequency, accounting for ~35% of the described mutations; these occur at codons 1061 and 1309 (Nagase and Nakamura, 1993; Miyaki *et al.*, 1994). Similarly there are two mutation hot spots in the somatic mutations of *APC*, clustered at codons 1309 and 1450 (Nagase and Nakamura, 1993).

### **1.2.3 The APC tumour suppressor protein; Structure and Protein Interactions**

The *APC* gene codes for a protein of 2844 amino acids (~310 kDa) containing a number of distinct domains and repeated elements (see figure 1.2) (Polakis, 1995; Polakis, 1997). Many of the repeat elements of the *APC* protein have been recognized as motifs involved in interaction with other proteins. To date a large number of these interactions have been confirmed and explored.

**Figure 1.2:** Schematic of some of the proteins central to the Wnt/Wg signalling pathway. Figures are not to scale.

**Figure 1.2: Protein structure of Wnt/WG pathway components**



Beginning with the amino terminal region of the protein the first recognized active region of the APC protein is an oligomerization domain, believed to contain >95%  $\alpha$ -helical conformation. It was established using immunoprecipitation experiments involving truncated and wild type proteins that APC has the ability to bind to itself (Su *et al.*, 1993). This domain appears to be formed of heptad amino-acid repeats, the most important of which are found in the interval amino acid 6-57 (Joslyn *et al.*, 1993; Su *et al.*, 1993). The ability of the APC protein to oligomerise may be responsible for the observations of dominant-negative effects of mutant APC proteins made by some researchers (Dihlmann *et al.*, 1999). Interestingly some of the alternative splice variants of the APC gene may fail to produce proteins containing the most important oligomerisation domain, presumably creating a protein with altered activity or regulation (Samowitz *et al.*, 1995). Additionally there is an attenuated form of FAP which appears to be characterised by the mutation of one allele of APC very near its 5' end, upstream of codon 157. The truncated protein produced in this syndrome seems to lead to a less severe FAP phenotype (fewer polyps with a reduced rate of cancer formation) while slightly longer truncated proteins seem to impart none of this reduction of the FAP phenotype (Grodén *et al.*, 1993; Spirio *et al.*, 1993; Spirio *et al.*, 1992). This attenuated FAP phenotype appears to arise when the initial, inherited APC mutation fails to prevent the APC gene product performing its normal cellular functions but makes the locus more easily inactivated by further mutations (Su *et al.*, 2000).

The next region as we move towards the carboxy terminus of the APC protein is a series of armadillo repeats (Polakis, 1997). These repeats are homologous to 42 amino acid motifs found initially in the *Drosophila* armadillo protein, the homologue of the human  $\beta$ -catenin protein (Hatzfeld, 1999). Multiple armadillo repeats appear to form a super-helical structure, containing a positively charged groove (Huber *et al.*, 1997). This segment of the APC protein contains 7 armadillo repeats and the function of these repeats appears to be protein-protein interaction, possibly needed for proper binding of  $\beta$ -catenin or for competitive binding to cadherins to the exclusion of  $\beta$ -catenin (Hirschl *et al.*, 1996). More recently, however, it has been shown that this set of repeated elements binds to

the regulatory domain of protein phosphatase 2a (PP2A), a protein capable of interacting with and involved in the regulation of Axin, another component of the Wnt-1 signalling pathway (Seeling *et al.*, 1999; Hsu *et al.*, 1999). These armadillo repeats are rarely lost due to mutations causing protein truncation and subsequent polyp and tumour formation suggesting that they may be N-terminal to those motifs which possess tumour suppressor function or that these armadillo repeats are absolutely vital to cell survival and mutants lacking them are selected against. In cases where mutation does cause a loss of some armadillo repeats from the APC protein the number of repeats lost shows a positive correlation with development of congenital hyperplasias of the retinas (Davies *et al.*, 1995; Wallis *et al.*, 1994; Wallis *et al.*, 1999). This may lend evidence to the observations by some researchers of a potential partial dominant-negative effect of truncated APC protein (Dihlmann *et al.*, 1999).

Within the segment of the APC protein containing the armadillo repeats, at Asp<sup>777</sup>, is a DNI/LD motif conserved across a number of species, which is recognized and cleaved by caspase-3 (Browne *et al.*, 1994; Webb *et al.*, 1999; Browne *et al.*, 1998). The caspases are final stage mediators of the apoptosis pathway, cleaving a number of target proteins, including poly-adenosylribose polymerase (PARP), DNA-dependant protein kinase (DNA-PK) and  $\beta$ -catenin (Woo *et al.*, 1998; Konopleva *et al.*, 1999; Casciola-Rosen *et al.*, 1996; Rosen and Casciola-Rosen, 1997; Webb *et al.*, 1999). The cleavage of APC during the process of apoptosis results in a 90 kDa amino-terminal fragment which contains a number of the armadillo repeat elements independent of the more carboxy terminal effector regions of APC.

Downstream of the armadillo repeat segment of the APC protein are found three 15 amino acid repeats, responsible for the binding of  $\beta$ -catenin (Su *et al.*, 1993). These repeats are not homologous to the motifs found on other proteins which interact with  $\beta$ -catenin (T-cell factor (TCF)/Lymphocyte enhancer factor (LEF) transcription factors and cadherins for example) and need not be removed by truncation mutation to eliminate the tumour suppressor functions of the APC

protein (Kintner, 1992; Ozawa *et al.*, 1995; Polakis, 1995; Beroud and Soussi, 1996).

Following the  $\beta$ -catenin binding repeats are a series of 7 relatively dispersed 20 amino acid motifs (Grodén *et al.*, 1991). These motifs are conserved within APC homologs in a number of different species including *Drosophila* (Hayashi *et al.*, 1997). These 20 amino acid repeats bear some resemblance to the 15 amino acid repeats found more N-terminal and were believed to play a role in the binding and subsequent degradation of  $\beta$ -catenin. This is in part supported by the fact that most tumourigenic truncation mutations of APC eliminate some or all of these motifs (Polakis, 1995). This may not be the whole story, however, as interspersed with these  $\beta$ -catenin binding and degradation domains are amino acid motifs which bind Axin (Zeng *et al.*, 1997). Axin is the gene product of the *fused* locus in the mouse, Axil/Conductin in *Xenopus* while the *Drosophila* homologue is called D-axin (Behrens *et al.*, 1998; Hamada *et al.*, 1999; Fagotto *et al.*, 1999). It appears that Axin acts as a scaffold protein in these diverse species, bringing together a number of proteins involved in the Wnt-1 signalling pathway (see section 1.3) in order to function as a regulator of free cellular  $\beta$ -catenin levels and hence regulate transcription of Wnt-1 target genes (Ikeda *et al.*, 1998; Kishida *et al.*, 1999; Sakanaka *et al.*, 1998; Sakanaka and Williams, 1999; Yamamoto *et al.*, 1999; Yamamoto *et al.*, 1998). It may be that the tumourigenic potential of the truncated APC proteins is not entirely due to the loss of the 20 amino acid repeats but rather due to the loss of Axin interaction elements preventing the formation of protein clusters necessary for  $\beta$ -catenin degradation.

Within the same region of the APC protein is one of three S(T)PXX motifs, the other two being located further C-terminal. These motifs have been recognised as being involved in protein interactions with A/T rich DNA sequences and it has been demonstrated that APC will co-immunoprecipitate genomic DNA with a high A/T content (Deka *et al.*, 1999). This may indicate a direct role for nuclear APC in the regulation of transcription or its involvement in the binding of proteins closely involved with transcription control. The decreasing gradient of nuclear APC along the crypt matches nuclear proliferation, lending further



support to a possible role in proliferation control (Midgley *et al.*, 1997; Smith *et al.*, 1993).

The next region of protein interaction on the APC protein as we move towards the C-terminal is the Basic Domain (Grodén *et al.*, 1991). This region, stretching approximately from amino acid 2200 to 2400, contains a higher proportion of the basic residues arginine and lysine, hence its nomenclature. This region of the APC protein enables it to bind to microtubules and full length APC is believed to be capable of promoting the polymerisation of tubulin *in vitro*, while mutant APC proteins are unable to encourage this reaction (Smith *et al.*, 1994; Munemitsu *et al.*, 1994). This ability to assemble and/or associate with microtubules may in part explain the localisation of APC to the leading edge of migrating cells, a region where microtubule assembly is required, as well as the observations of disordered cell migration in models of APC overexpression and in animal models of FAP (Nathke *et al.*, 1996; Mahmoud *et al.*, 1999; Mahmoud *et al.*, 1997; Wong *et al.*, 1996).

In addition to interacting with A/T rich DNA sequences, the C-terminal third of APC also interacts with at least two other proteins; EB1 and Human Discs Large (hdlg). EB1 has been shown to bind in the final 284 amino acids of APC while hdlg binds in the final 72 amino acids (Su *et al.*, 1995; Matsumine *et al.*, 1996). EB1 was found to bear little resemblance to any protein of known function when it was first isolated, although its coding sequence was similar to one yeast and two human sequences. More recently EB1 has been found to be homologous to *Mal3*, a yeast protein found to associate with microtubules and a number of cellular structures involved in mitosis, including centrosomes and spindle poles (Beinhauer *et al.*, 1997; Morrison *et al.*, 1998). EB1 therefore may be partially responsible for the observation of APC localisation to these cell division structures and for its close association with microtubules. The hdlg protein is the human homolog of the *Drosophila* Discs Large (dlg) tumour suppressor protein; loss of dlg in *Drosophila* results in neoplastic growth in the imaginal disc of the developing embryo (Lue *et al.*, 1994; Woodhouse *et al.*, 1998). APC interacts with hdlg/dlg through conserved S/TXV domains in its C-terminal portion, which have been found to bind to the PDZ domain of the discs large homologs.



Recently, in addition to its innate roles in cell cycle control hdlg has been found to play a role in APC induced control of the cell cycle in mammalian systems, implying that loss of the terminal regions of the APC protein may result in neoplastic transformations through modulation of hdlg activity as well as through mediation of  $\beta$ -catenin levels (Ishidate *et al.*, 2000).

Also in the C-terminal third of the APC protein are a number of target sites for the Cyclin Dependant Protein Kinase p34<sup>cdc2</sup>. Eleven potential phosphorylation sites {(S/T)PX(R/K)} have been identified in the region of APC most often deleted in truncation mutation. (Trzepacz *et al.*, 1997) These sites appear to be phosphorylated in a cell cycle dependant manner, with the highest levels of phosphorylation occurring during mitosis (Bhattacharjee *et al.*, 1996). Alteration of the phosphorylation state of APC appears to alter its binding activities and its ability to down-regulate levels of  $\beta$ -catenin.

#### **1.2.4 Expression pattern**

The *APC* gene is expressed in many human and animal epithelial tissues, the gene products including a number of protein isoforms which appear to be restricted to a small number of tissues, the most apparent being the Brain Specific isoform which is encoded using an alternative first exon (Midgley *et al.*, 1997; Bardos *et al.*, 1997). The expression of *APC* in epithelial tissues seems to be strongest in differentiated cells (Midgley *et al.*, 1997). Immunoelectronmicroscopy reveals that the *APC* gene product localises in both the nucleus and the cytoplasm of mammalian cells (Neufeld and White, 1997; Miyashiro *et al.*, 1995). In the former the protein seems to cluster into distinct regions, possibly nucleoli. In the latter, the protein appears in some cases to cluster on the ends of microtubules, particularly at the leading edge of actively migrating cells, hinting at its role in cellular adhesion and morphology maintenance (Morrison *et al.*, 1997; Nathke *et al.*, 1996). The cytoplasmic staining patterns vary from tissue to tissue, in some cases showing staining predominantly apically, in others showing strong staining on the lateral borders of cells (Midgley *et al.*, 1997). Levels of *APC* expression appear to be variable

and may depend on a number of factors, including developmental signals as in differentiating rat neuronal cells (Bhat *et al.*, 1994; Dobashi *et al.*, 1996; Jaiswal and Narayan, 1998).

### 1.2.5 Function

The role of APC in the prevention of tumour development, or more accurately the failure of truncated APC to prevent neoplastic transformation, has led to an intense study of its role in normal and dysregulated cell functioning. Like many tumour suppressor proteins APC appears to play a role in the embryonic development of organisms as well as in the maintenance of health in adult animals. Homologs of APC appear to play important roles in pattern development in a number of diverse species. In *Drosophila* an APC equivalent acts as a signalling step in the wingless pathway (Hayashi *et al.*, 1997). *C. Elegans* also utilises an APC homologue in the *Wnt1*/wingless pathway; in embryonic development this pathway functions to promote differentiation of cells into endoderm (Rocheleau *et al.*, 1997). In *Xenopus* an overexpression of another APC-related gene brought on by injected RNA results in the formation of a second dorsoanterior axis and Wnt signalling appears to be involved in the formation of the Spemann organiser (Vleminckx *et al.*, 1997; Laurent *et al.*, 1997). Studies in the rat have revealed that the APC gene is upregulated in the developing nervous system, particularly during the differentiation of certain cell types (Bhat *et al.*, 1994; Dobashi *et al.*, 1996). It would appear that APC and its homologues are vital in a number of developmental processes, including cellular differentiation, segment polarity and axis induction. Studies of murine Apc's role in development also indicate a vital embryonic dependence on its proper functioning (see section 1.4.2) (Moser *et al.*, 1995).

It is the involvement of APC and its many cross species homologs in the *Wnt*/wingless signalling pathway which appears to be primarily responsible for its influence on cellular growth and differentiation, both in healthy and cancerous tissues. It is easiest, perhaps to elucidate the pivotal role of APC in this pathway by outlining the functions of the other signalling proteins involved in this cascade.

### **1.3 The Wnt/Wingless signalling pathway**

#### **1.3.1 Components of the Wnt Pathway**

##### **1.3.1.1 Wnt**

The Wnt proteins are a family of secreted signalling molecules with a number of actions in a variety of organisms. The Wnt family was first identified in the mouse, when two family members (initially named *int1* and *2*) were discovered to behave as oncogenes when activated by the insertion of Mouse Mammary Tumour Virus (MMTV) in mammary tumours (Nusse and Varmus, 1982). This correlation with tumour development lead to a great deal of study, *Wnt-1* homologs were quickly identified in a wide range of species including human, *C. elegans* and *Drosophila*; in the latter the conserved protein was known as *wingless* (*wg*) (Cabrera *et al.*, 1987). In all of these species the Wnt/wingless family of proteins play important developmental roles; defining segment polarity in *Drosophila*, controlling gut development in *C. elegans* and determining axis formation in *Xenopus* (Nusslein-Volhard, 1991; Nusslein-Volhard and Roth, 1989; Han, 1997; McMahon and Moon, 1989a; McMahon and Moon, 1989b). Subsequently, a large number of Wnt proteins (see <http://www.stanford.edu/~rnusee/wntwindow.html> and Nusse, 1999) have been identified in the mouse, all of which also have important developmental role and a number of which appear to be able to play a part in the development of cancers, particularly mammary tumours (Mester *et al.*, 1987; Nusse *et al.*, 1990; Tsukamoto *et al.*, 1988).

##### **1.3.1.2 Frizzled**

Frizzled or frizzled homologs have been found to be the receptor for the secreted Wnt signalling family (see figure 1.3) (Vinson and Adler, 1987; Zheng *et al.*, 1995; Yang-Snyder *et al.*, 1996). The frizzled loci of a wide range of species encode a seven pass transmembrane protein; the cysteine-rich N-terminal portion being exterior with the C-terminal portion extending into the cytoplasm (Vinson *et al.*, 1989; Bhanot *et al.*, 1996). This exterior N-terminal region, containing 10 highly conserved cysteine residues is believed to interact with the circulating

Wnt proteins and thereby transfer growth signals (Bhanot *et al.*, 1996). It is not completely clear how the interactions of some Wnt and Frizzled pairings can lead to downstream modulation of the rest of the Wnt pathway. The C-terminal region of many of the Frizzled molecules contain a S/TXV motif, a candidate for binding with the PDZ domain (Songyang *et al.*, 1997). Dishevelled contains one of these domains at its C-terminus but evidence to date suggests that it does not interact directly with Frizzled (Cadigan and Nusse, 1997; Nusse, 1997; Nusse *et al.*, 1997). There is some suggestion that G-protein signalling may be involved. Some Wnt/Frizzled effects, such as regulation of Protein Kinase C in *Xenopus*, can be blocked through the application of G-protein inhibitors while two of the target proteins in the Wnt/wingless cascade, Axin and Dishevelled, contain motifs found in other molecules known to be involved in G-Protein signalling (Slusarski *et al.*, 1997; Fagotto *et al.*, 1999; Sheldahl *et al.*, 1999).

#### 1.3.1.3 Dishevelled

The next group of molecules involved in the transduction of the Wnt signal are the dishevelled (dsh) protein and its cross-species homologs (see figure 1.2). First characterised in *Drosophila*, dsh, a protein of approximately 650 amino acids, has been determined to be vital to the conduction of developmental *Wingless* signals (Klingensmith and Nusse, 1994a; Klingensmith *et al.*, 1994b; Noordermeer *et al.*, 1994; Theisen *et al.*, 1994). The dsh protein appears to fall between *frizzled* activation and modulation of the next step in the Wg cascade; zesty-white (the homolog of the mouse glycogen synthase kinase-3 $\beta$  or GSK-3 $\beta$ ) (Siegfried *et al.*, 1992; Noordermeer *et al.*, 1994; Dominguez *et al.*, 1995; Butz and Larue, 1995; He *et al.*, 1995). In the mouse, there are at least 2 homologs of the dsh protein, dvl-1 and dvl-2, this multiplicity may explain why the dvl-1 knock-out mouse develops normally, showing only subtle neurological and behavioural phenotype (Lee *et al.*, 1999; Lijam *et al.*, 1997). These proteins exhibit very similar properties to *Drosophila* dsh, including co-immunoprecipitation with casein kinase II (CKII), hyperphosphorylation on overexpression of themselves or of frizzled homologs and the ability to suppress GSK-3 $\beta$  activity (Ruel *et al.*, 1999; Lee *et al.*, 1999). It is not entirely clear how the Wnt/Wg signal is conveyed through dsh proteins; in response to Wnt/Wg

cellular dsh becomes hyperphosphorylated on serine and threonine residues and tends to associate more closely with the plasma membrane (Yanagawa *et al.*, 1995; Axelrod *et al.*, 1998; Klingensmith *et al.*, 1996). Proteins of the *dishevelled* family contain at least 4 conserved domains: there are 50 amino acids referred to as the DIX domain in the N-terminal portion of the protein which resemble a region of *axin* (another of the proteins involved in the Wnt/Wg pathway) as well as conserved PDZ and DEP domains in the central and C-terminal regions (see figure 1.2) (Vasicek *et al.*, 1997; Zeng *et al.*, 1997; Klingensmith *et al.*, 1994b; Klingensmith *et al.*, 1996; Heslip *et al.*, 1997; Theisen *et al.*, 1994; Ponting *et al.*, 1997). The *dishevelled* family of proteins also appear to play a role in another signalling pathway; one of its conserved domains, the DEP-domain, is implicated in the stimulation of the c-Jun/JNK pathway (Li *et al.*, 1999). Additionally, there is some evidence that, at least in *Drosophila*, *dsh* proteins can interact with and influence the activity of the important developmental protein *notch* (Rulifson *et al.*, 1996).

#### 1.3.1.4 Glycogen Synthase Kinase 3 $\beta$ (GSK-3 $\beta$ )

GSK-3 $\beta$  and its *Drosophila* homolog zeste-white 3 (zw3) are cytoplasmic serine/threonine protein kinases with a number of target proteins in the Wnt/Wg pathway (Ruel *et al.*, 1999). Targets of GSK-3 $\beta$  and zw3 include  $\beta$ -catenin, axin and APC as well as other related proteins such as the microtubule binding protein Tau (Rubinfeld *et al.*, 1996; Wagner *et al.*, 1996). The primary role of GSK-3 $\beta$  and its homologs in the conduction of the Wnt/Wg pathway appears to be in the phosphorylation of  $\beta$ -catenin, a reaction that ultimately marks it for degradation (see figure 1.3) (Aberle *et al.*, 1997). The current model is that Wnt/Wg signalling acts through frizzled and dishevelled homologs to downregulate the activity of GSK-3 $\beta$ , thereby reducing the phosphorylation of the target proteins involved in the Wnt/Wg pathway and creating an increase in the free cytoplasmic pool of  $\beta$ -catenin through reduction in its rate of turnover. GSK-3 $\beta$  also appears to bind other proteins which may influence its activity and may play some role in regulation of the Wnt/Wg pathway; one of these proteins is Frequently Rearranged in Advanced T-cell lymphomas, or FRAT1 (Thomas *et al.*, 1999).

This peptide appears to bind GSK-3 $\beta$  in such a way as to inhibit its interaction with Wnt cascade proteins but does not inhibit its binding of other targets such as glycogen synthase.

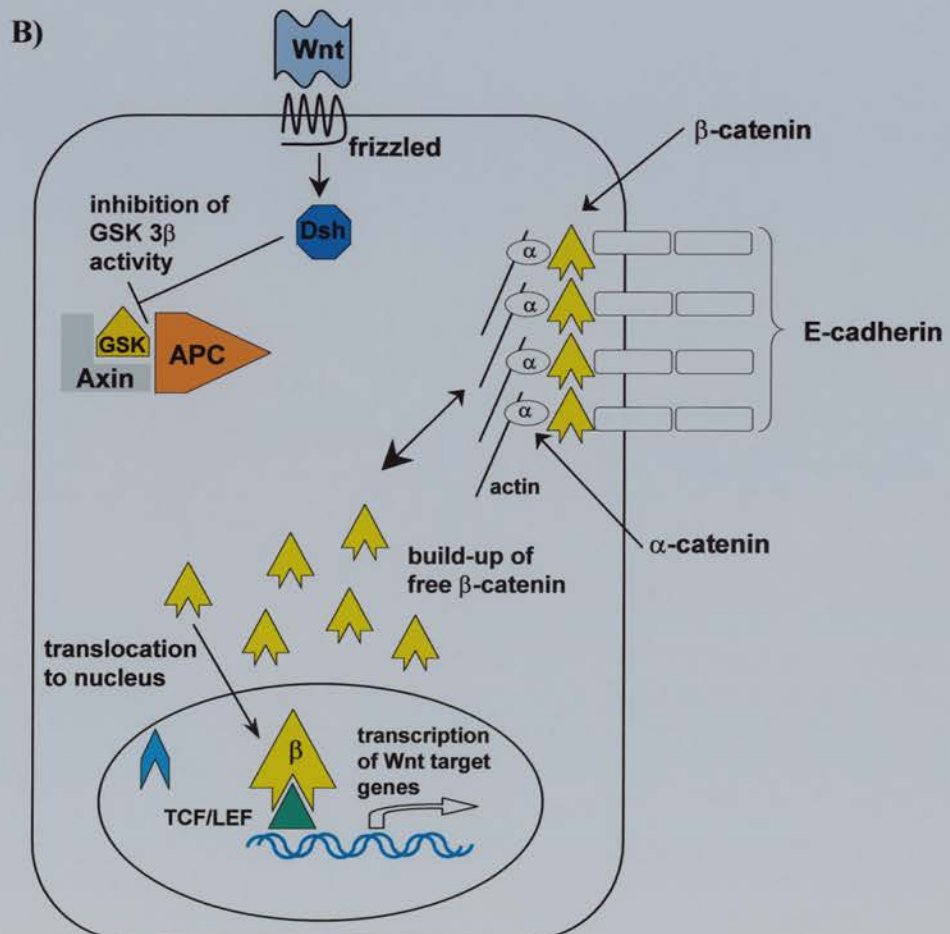
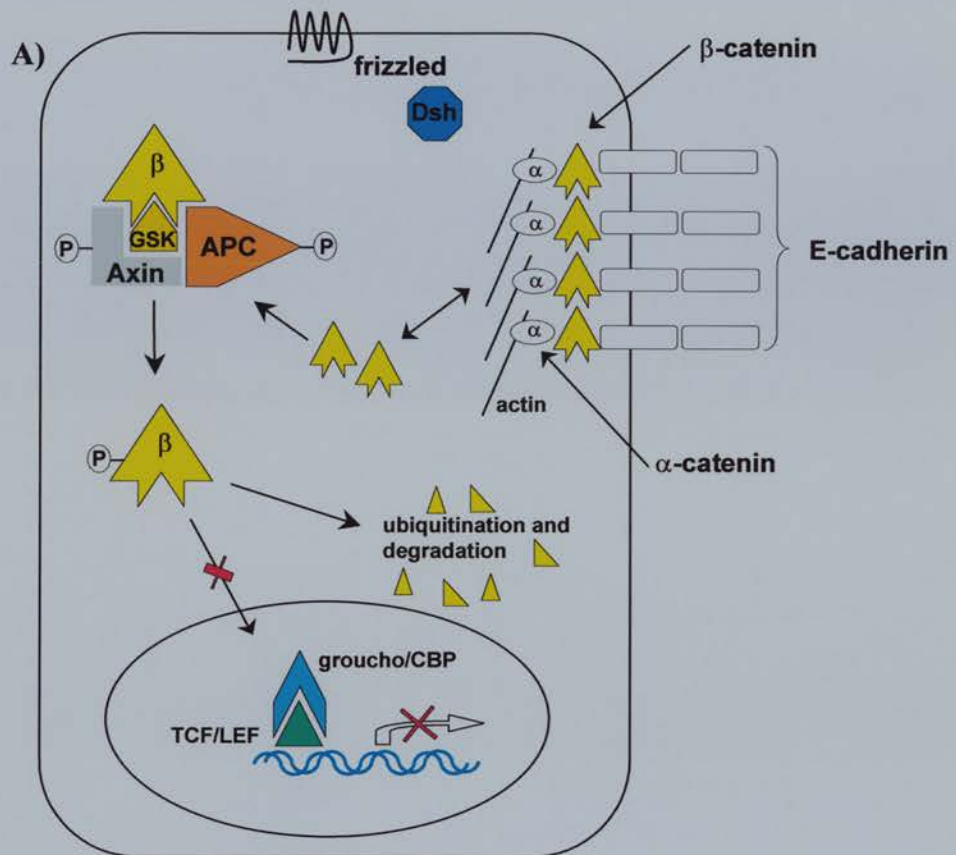
#### 1.3.1.5 Axin

While  $\beta$ -catenin was known to be a phosphorylation target of GSK-3 $\beta$ , it was discovered that, *in vitro*, this reaction is very inefficient (Yost *et al.*, 1996). The suggestion was that another cellular factor must be involved, increasing the efficiency of the reaction. A candidate was soon identified in the mouse, Axin, the gene product of the *fused* locus, so named because mutation at this locus results in axial duplication (Zeng *et al.*, 1997). As mentioned previously, Axin bears a resemblance in its N-terminal region to the *dvl-1* protein and also to G protein signalling (RGS) proteins (Zeng *et al.*, 1997). The role of the RGS region of Axin is not clear, although its loss through mutation seems to lead to dominant negative effects in *Xenopus*, resulting in axis duplication as seen in the *fused* mouse (Vasicek *et al.*, 1997). The roles of other regions of Axin are much better understood, these regions include motifs for the binding of a number of different proteins including the armadillo repeat region of  $\beta$ -catenin, GSK-3 $\beta$ , APC and itself through a C-terminal DIX dimerisation domain (see figure 1.2) (Behrens *et al.*, 1998; Nakamura *et al.*, 1998; Hamada *et al.*, 1999; Fagotto *et al.*, 1999; Sakanaka *et al.*, 1998; Sakanaka and Williams, 1999; Ikeda *et al.*, 1998; Hart *et al.*, 1998; Zeng *et al.*, 1997). There is strong evidence to suggest that axin acts as a scaffold, bringing together a number of different proteins to induce phosphorylation of  $\beta$ -catenin by GSK-3 $\beta$  with the participation of APC (see figure 1.3) ultimately resulting in the degradation of  $\beta$ -catenin (Behrens *et al.*, 1998; Hamada *et al.*, 1999; Nakamura *et al.*, 1998).

**Figure 1.3:** A simplified diagram of the Wnt pathway signalling cascade. **A)** In the absence of Wnt binding to membrane bound frizzled proteins the GSK-3 $\beta$ /Axin/APC complex retains phosphorylation and promotes the phosphorylation and subsequent ubiquitination/degradation of  $\beta$ -catenin. Wnt target genes remain repressed by Groucho/CBP binding of the TCF/LEF transcription factors. **B)** On Wnt binding to the surface frizzled proteins *dishevelled* represses the activity of the GSK-3 $\beta$ /Axin/APC complex and  $\beta$ -catenin accumulates in the cytoplasm and can translocate to the nucleus where it displaces Groucho/CBP to activate transcription through interaction with TCF/LEF.



**Figure 1.3: The Wnt/Wg signalling pathway**





### 1.3.1.6 $\beta$ -catenin

$\beta$ -catenin plays at least two important roles in the cell, firstly mediating cellular adhesion through its interactions with E-cadherin and, indirectly, the cytoskeleton, and secondly functioning as the lynchpin in the control of expression through the Wnt/Wingless pathway (Gumbiner and McCrea, 1993; Butz and Kemler, 1994; Behrens, 1999).

#### 1.3.1.6.1 Structure and function

$\beta$ -catenin shows a great degree of cross-species conservation both in its structure and its function. Human  $\beta$ -catenin consists of 781 amino acids arranged broadly into 3 important domains (see figure 1.2) (Hulsken *et al.*, 1994a). The first comprises the N-terminal 130 amino acids and contains a number of phosphorylation target residues for the GSK-3 $\beta$  (Yost *et al.*, 1996). The central section of  $\beta$ -catenin contains twelve 42 amino acid armadillo repeats and forms an important protein interaction domain while the carboxy terminal 100 amino acids form a region of interaction with nuclear transcription factors (Huber *et al.*, 1997; van de Wetering *et al.*, 1997; Behrens *et al.*, 1996; Vleminckx *et al.*, 1999).

The most N-terminal domain, the initial 130 amino acids, contains a number of conserved serine and threonine residues (Polakis, 1999; Rubinfeld *et al.*, 1997; Orford *et al.*, 1997). These residues have been found to be phosphorylated *in vivo* and at least one kinase, GSK-3 $\beta$  in co-operation with the scaffold protein Axin, appears to be responsible (Yost *et al.*, 1996; Zeng *et al.*, 1997). As a clue to the function of this phosphorylation it has been found that N-terminal truncation mutations of  $\beta$ -catenins artificially introduced into cells tend to accumulate in the cytoplasm, escaping the normal turn-over mechanisms (Barth *et al.*, 1997; Munemitsu *et al.*, 1996). Additionally, in many tissues, activation mutations of  $\beta$ -catenin, those that result in its cellular accumulation and oncogenic effects, often involve those bases coding for the N-terminal serine and/or threonine residues (Iwao *et al.*, 1998; Miyoshi *et al.*, 1998a; Miyoshi *et al.*, 1998b; de La

Coste *et al.*, 1998; Fukuchi *et al.*, 1998). The residues involved in most of the activation mutations have been found to be serines 33, 37 and 45 as well as threonine 41 (Morin *et al.*, 1997; Polakis, 1999). The phosphorylation of these residues has been hypothesised to regulate the stability of free pools of  $\beta$ -catenin in the cell by modulating the targeting of  $\beta$ -catenin to the ubiquitination/proteasome pathway (Aberle *et al.*, 1997; Orford *et al.*, 1997). The importance of the maintenance of appropriate concentrations of free  $\beta$ -catenin in the cytoplasm to normal cell function is highlighted by the apparent mutual exclusiveness of  $\beta$ -catenin C-terminal mutations and APC truncation mutations in the same neoplasias; nearly all colorectal cell lines as well as spontaneous or induced tumours of the colon show either one or the other mutation, but not both (Ilyas *et al.*, 1997; Sparks *et al.*, 1998). With the majority of phenotypic mutations of APC leading to a lack of  $\beta$ -catenin binding and phosphorylation and the majority of  $\beta$ -catenin mutations resulting in phosphorylation incompetence, this suggests that colorectal carcinogenesis may have a near absolute requirement for interruption of normal  $\beta$ -catenin phosphorylation through APC interaction. One mutation in this pair of proteins may lead to a reduction in selective pressure for another mutation.

The central portion of the  $\beta$ -catenin protein, as mentioned previously, consists of 12 armadillo repeats. As previously mentioned armadillo repeats together form a complex 3-dimensional structure, encompassing a tightly packed series of short  $\alpha$ -helices with a positively charged central groove, usually associated with protein-protein interactions (Huber *et al.*, 1997). Indeed, this central region does appear to be involved in the interaction of  $\beta$ -catenin with a number of other proteins, including T-cell Transcription Factor 1 (TCF-1)/Lymphocyte Enhancer-binding Factor 1 (LEF-1), APC, E-cadherin, Axin and the actin associate protein Fascin (Molenaar *et al.*, 1996; Behrens *et al.*, 1996; Rubinfeld *et al.*, 1995; Hulsken *et al.*, 1994b; Nakamura *et al.*, 1998; Tao *et al.*, 1996). Many of these interactions appear to occur in overlapping portions of the armadillo tertiary structure groove; indeed, some of these protein interactions appear to be mutually exclusive (Hulsken *et al.*, 1994b). The evidence for the interaction of these proteins with this portion of the  $\beta$ -catenin molecule is supported by the fact that

many of these proteins are acidic and as such are negatively charged at cellular pH. Additionally, the phosphorylation of APC by GSK 3 $\beta$ , which increases the stability of  $\beta$ -catenin/APC interactions, also increases the negative charge of APC (Huber *et al.*, 1997). It is this central section of  $\beta$ -catenin which is perhaps most responsible for its cellular effector functions. The interactions of the  $\beta$ -catenin armadillo repeat domain with E-cadherin and with  $\alpha$ -catenin play a role in cellular adhesion; E-cadherin is a calcium ion dependant cell adhesion molecule, whose cytoplasmic domains bind either  $\beta$ -catenin or  $\gamma$ -catenin (plakoglobin) (Aberle *et al.*, 1996). Either catenin molecule ( $\gamma$  or  $\beta$ ) is then complexed with  $\alpha$ -catenin which in turn acts as a bridge to bind to actin molecules making up the cytoskeleton (Aberle *et al.*, 1996). In this way  $\beta$ -catenin links cell surface adhesion molecules to the cell architecture.

The carboxy terminal region of  $\beta$ -catenin has been determined to act as a transactivation domain, influencing transcription of target genes when binding TCF/LEF transcription factors (van de Wetering *et al.*, 1997; Vlemminckx *et al.*, 1999).  $\beta$ -catenin appears to displace the corepressor molecule groucho from these High Mobility Group (HMG) transcription factors when cytoplasmic levels of free  $\beta$ -catenin have reached a sufficient concentration (see figure 1.3) (Roose *et al.*, 1998). An additional repressor of this transcriptional activation appears to be CREB-binding protein (CBP) which also binds TCF transcription factors (Waltzer and Bienz, 1998).

#### **1.3.1.6.2 Downstream targets of $\beta$ -catenin**

As mentioned above,  $\beta$ -catenin associates with the HMG transcription factors LEF-1 and the TCF family, especially TCF-4, and contains a transactivation domain (Aoki *et al.*, 1999; Clevers and van de Wetering, 1997; van de Wetering *et al.*, 1997). In order to have an effect on levels of transcription of target genes the  $\beta$ -catenin/transcription factor complex must locate to the nucleus. Interestingly it appears that  $\beta$ -catenin can translocate into the nucleus independent of importins and without containing a nuclear localisation signal (Yokoya *et al.*, 1999; Fagotto *et al.*, 1998). Additionally it appears that binding

of  $\beta$ -catenin to the LEF transcription factors can increase the uptake of the complex by the nucleus (Huber *et al.*, 1996b). Although the mechanism for  $\beta$ -catenin/transcription factor nuclear localisation is unclear at present, it is clear that the complex upregulates transcription levels of a number of target genes. Genes which have been identified as targets of the  $\beta$ -catenin/LEF/TCF transcriptional activation include a number of genes thought to be involved in adhesion or growth control and neoplasia; these include c-myc, cyclooxygenase-2, cyclin D1, fibronectin and E-cadherin among others (He *et al.*, 1998; Howe *et al.*, 1999; Shtutman *et al.*, 1999; Gradl *et al.*, 1999; Huber *et al.*, 1996a). Other genes which have been implicated in neoplastic growth have also been found to be targets of  $\beta$ -catenin/TCF activation, an example being the matrilysin metalloproteinase, which appears to confer growth potential to cells and Cdx1, a murine homeobox gene which is downregulated in most human colorectal cell lines (Brabletz *et al.*, 1999; Wilson *et al.*, 1997; Crawford *et al.*, 1999; Lickert *et al.*, 2000). The number of genes identified as targets of the  $\beta$ -catenin/TCF complex is rapidly increasing and encompass genes involved in a large number of cellular pathways (see <http://www.stanford.edu/~rnusse/wntwindow.html>). The final effects of Wnt activation would appear to be very complex on the level of gene transcription control, although the final effects on development and neoplasia are perhaps better understood.

### 1.3.2 The Wnt pathway in development

It appears that the Wnt pathway plays a very important role in pattern development in a number of diverse species. In *Drosophila*, *C. Elegans*, *Xenopus* and the mouse the wingless pathway seems to be intimately involved in the specification of cellular differentiation and/or in the development of body axis (Hayashi *et al.*, 1997; Rocheleau *et al.*, 1997; Vleminckx *et al.*, 1997; Dobashi *et al.*, 1996). In *Drosophila* the Wingless pathway is essential to the induction of segment polarity, while in *C. Elegans* this pathway functions to promote differentiation of cells into endoderm (Cabrera *et al.*, 1987; Rocheleau *et al.*, 1997). In *Xenopus* Wnt/Wingless signalling appears to control the formation of dorsoanterior axis; injections of a number of Wnt RNAs which result in the overexpression of various members of the Wnt family has been observed to result in the duplication of this axis (Wodarz and Nusse, 1998; Miller and Moon, 1997; Fagotto *et al.*, 1997; Itoh *et al.*, 1998). Also in *Xenopus* the Wnt pathway appears to interact with the SMAD/STAT pathway in the assembly of the Spemann organiser (Laurent *et al.*, 1997; McKendry *et al.*, 1997). Additionally it would appear that Wnt and its homologues are vital to the process of cellular differentiation. Studies in the rat have revealed that the *APC* gene is upregulated in the developing nervous system, particularly during the differentiation of certain cell types (Bhat *et al.*, 1994; Dobashi *et al.*, 1996).

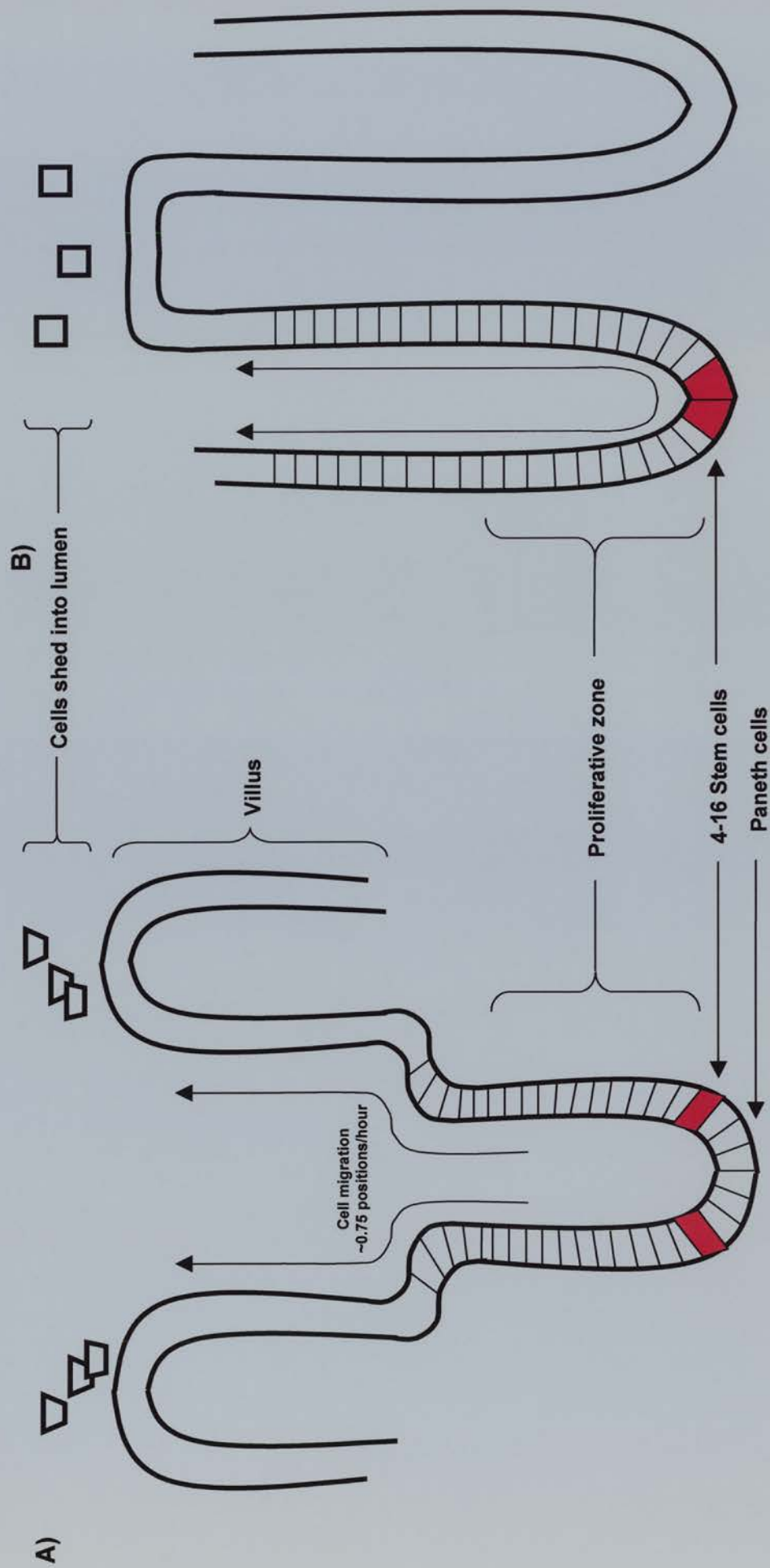
This process of differentiation takes place not only during embryogenesis, during which structures and organs are produced, but also occurs in certain tissues of the mammal in a life-long manner. Some examples include the haemopoietic system, wherein mature blood cells differentiate from progenitor cells (stem cells) and the epithelium of the intestine, where cells are continuously sloughed into the lumen to be replaced by cells derived from stem cells found in intestinal crypt structures. In the mouse the development of the intestine begins on the seventh day of gestation, with connection to the oral cavity occurring on day 9 and fusion with the hindgut to give rise to the anal opening on day 10 (Gilbert, 1994). While the large-scale structures of the mouse gut are in place prenatally, the maturation of the epithelial lining is not complete until some weeks after birth. Intestinal crypts, the basic unit of cell proliferation in the mammalian large and

small intestine are fully mature until between two and three weeks postnatally (Calvert and Pothier, 1990). The crypts are, in a sense, a source of cells in continuous development. At the base of the crypts are found a number of stem cells, which divide continuously to give daughter cells which are themselves capable of division although not indefinitely. The number of stem cells and their location differ in the large and small intestine (Potten and Booth, 1997). In the small intestinal crypts the 4-16 stem cells, capable of division every ~24 hours, are found arranged in a ring at the fourth position from the base of the crypt which is occupied by differentiated paneth cells (see figure 1.4a). In the large intestine a similar number of stem cells are found in the very base of the crypt (see figure 1.4b) (Bach *et al.*, 2000). The process of cell division in the crypt is coupled with a continuous migration of these cells out from the base of the crypt to the villi of the small intestine or the luminal surface in the large intestine where the cells are eventually sloughed off. During this migration the cells become less and less capable of division and become more differentiated. This process gives a crypt to lumen gradient of differentiation, marked by changing expression patterns including expression of many of the proteins involved in the Wnt signalling pathway including  $\beta$ -catenin and APC (Nathke *et al.*, 1996; Nathke *et al.*, 1994). This constant process of differentiation may be driven in part by the Wnt pathway and may explain the central importance of this pathway in the maintenance of growth control and prevention of carcinogenesis in the mammalian intestine.

**Figure 1.4:** **A)** The basic structure of the small intestinal crypt. The crypt is composed of ~250 cells of which 4-16 are stem cells and some 150 retain proliferative capability. The cells migrate out of the crypt and onto adjacent villi from which they are eventually shed into the intestinal lumen. **B)** The large intestinal crypt differs in being somewhat larger with its stem cells occupying the lowest position.



Figure 1.4: Crypt structure





### 1.3.3 Dysregulation of the Wnt /Wingless pathway

The role of the Wnt pathway in the control of embryonic development and the differentiation of cells in particular tissues has been largely elucidated by study of its dysregulation. Axis duplication and failure of segment polarity development are just some of the disruptions of development possible with altered Wnt/Wingless signalling. Additionally, disruption of normal Wnt signalling in the adult organism can result in a disease phenotype, most notably in the development of neoplastic disorders. Neoplastic cells are those which have gained the ability to grow independent of physiological regulation; this regulation may take the form of internal signals such as those which promote apoptosis or external signals such as growth factors or contact inhibition. While the events leading to this loss of control are not yet fully understood it has become clear that such a dramatic change in cellular behaviour is not the result of a single factor but more often the result of accumulated genetic alterations (Vogelstein and Kinzler, 1993). The “multi-step” progression from normal to neoplastic cell has been well studied in colorectal cancers, where mutations in genes involved in the Wnt pathway have been found to be common events. The *APC* gene has been found to be mutated in over 80% of both sporadic adenomas and sporadic human colorectal carcinomas and so is strongly linked to tumour development and  $\beta$ -catenin has also been found to be mutated in a significant proportion of these lesions (Smith *et al.*, 1993; Jen *et al.*, 1994; Iwao *et al.*, 1998). Study of these colorectal tumours has also suggested that the order in which genetic lesions occur is integral to tumour formation: in the case of colorectal carcinoma the loss of wild-type APC function is thought to be a very early event (Luongo *et al.*, 1994; Powell *et al.*, 1992). The sequence of events in the genesis of a colorectal tumour is thought to begin with the loss of the “gatekeeper” function of APC, in which APC is thought to work to integrate several cellular signals to limit cell proliferation. Loss of this function leads to altered cell proliferation; colonic crypts become dysplastic, accumulating cells more rapidly than they are lost or absorbed thus leading to polyp formation. One or more of these accumulating cells may then go on to experience mutations in cellular oncogenes, such as *RAS* which allows more prolific cell growth. The

rapid growth of these cells, possibly coupled to deficiencies in DNA repair mechanisms, may eventually lead to the loss of a number of tumour suppressor genes (such as *p53*) allowing the more rapid accumulation of genetic damage. These altered cells may finally accumulate further changes which allows them to become invasive, achieving metastasis.

As mentioned above it is not only mutations in *APC* which can result in Wnt pathway mediated growth control escape. The first identified Wnt gene was found as a result of its virus integration activation leading to breast cancer in mice (Nusse and Varmus, 1982; Li *et al.*, 2000).  $\beta$ -catenin mutations have been detected in colorectal tumours, as well as in hepatocellular carcinomas and ovarian carcinomas (de La Coste *et al.*, 1998; Gamallo *et al.*, 1999). Additionally levels of a secreted frizzled homolog appear to be depressed in some breast cancers, possibly enhancing regular Wnt signalling interactions by freeing more Wnt ligand to bind with membrane bound frizzled molecules (Zhou *et al.*, 1998). The implication of many of the members of the Wnt pathway in various neoplastic disorders, primarily deriving from epithelial tissues, suggests a fundamental role for this pathway in the control of growth and differentiation in this tissue type.

It has been relatively recently discovered that the Wnt pathway is modulated *in vitro* in response to genotoxic insult, leading to speculation that this regulation may also occur *in vivo*. Human derived colorectal cell lines as well as murine embryonic fibroblasts appear to upregulate levels of both *APC* mRNA and protein in response to treatment with N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), a powerful methylating agent (Jaiswal and Narayan, 1998). This *APC* upregulation may lead to the increased degradation of  $\beta$ -catenin and concomitant downregulation of Wnt target genes. Interestingly, the methylation dependant induction of *APC* appears to be dependant on functional *p53* tumour suppressor protein activity as cell lines and primary cultures lacking *p53* failed to show a response when subjected to MNNG (Narayan and Jaiswal, 1997).

## 1.4 The *Min* Mouse; a murine model of FAP

The human syndrome FAP (section 1.1.2.2) led to the discovery and characterisation of the human *APC* gene and its product. The role of *APC* as a potent tumour suppressor in the intestine, and to a lesser degree in other tissues, suggested that an understanding of the cellular functions of *APC* would lead to a more global understanding of the process of tumour formation. The in depth study of *APC*'s interactions and the changes in cellular responses following *APC* mutation would be near impossible if human tumours or tissues from FAP patients were the only sources of experimental models. What was needed was an animal model.

Random mutagenesis using the point mutagen Ethylnitrosourea (ENU) created a mutant which displayed two unusual phenotypes caused by separate mutations; it tended to run in circles and it rapidly developed anaemia (Moser *et al.*, 1990). The anaemia was soon found to be caused by bleeding from the intestine, itself precipitated by the presence of a number of polyps; the mice were termed *Min* (multiple intestinal neoplasia) mice and the phenotype was eventually traced to a heterozygous truncation mutation at codon 850 in the murine homolog of the human *APC* gene (Moser *et al.*, 1990; Su *et al.*, 1992). Because of the obvious parallels between this mutant mouse line and the human FAP patients, the *Min* mice soon became involved in the study of tumour development.

### 1.4.1 Phenotype of the *Min* Mouse

While mainly phenotypically normal the *Min* mouse, like the human FAP patient, is prone to development of numerous intestinal polyps which have lost the remaining wild type copy of murine *Apc*, eventually leading to tumour development (Luongo *et al.*, 1994). Interestingly the polyps and subsequent tumours in the *Min* mouse tend to develop primarily in the small intestine, while in the human they are largely associated with the large intestine and colon (Moser *et al.*, 1995; Chow *et al.*, 1996; Neugut *et al.*, 1998). While the *Min* phenotype is eventually lethal it has nevertheless enabled the study of many aspects of the genetics of intestinal cancers (Moser *et al.*, 1990). Once the *Min*

mouse line was phenotypically characterised it was used in the search for factors which may modulate the penetrance of the *Apc* mutation initiated adenocarcinomas. Matings with mouse strains bearing other genetic alterations, provided clues to the mediators of *Apc* function and its role in neoplasia. One of the first genes which was discovered to alter the severity of the *Min* phenotype was *Mom-1* (Modulator of *Min*), which was mapped to the distal portion of chromosome 4 (Dietrich *et al.*, 1993). It was discovered that the *Mom-1* background on which the *Min* genotype was crossed could alter the number of intestinal polyps formed by several fold and that ~50% of this change could be directly attributed to the genetic influence of *Mom-1* (Dietrich *et al.*, 1993). Recently the *Mom-1* gene product has been identified as the secretory phospholipase-a2 (*Pla2g2*) which is expressed in the murine small intestine and is involved in the cleavage of fatty acids from lipids (MacPhee *et al.*, 1995; Gould *et al.*, 1996a; Gould *et al.*, 1996b) Although it is yet to be understood how this molecule modulates the *Min* phenotype it is known that all mouse strains displaying sensitivity to the *Min* phenotype contain a premature stop codon in the *Pla2g2* gene and *Min* sensitive mouse lines which have been engineered to express a full length *Pla2g2* transgene construct which does not contain the premature stop codon display markedly reduced polyp formation (Cormier *et al.*, 1997).

Mice bearing the *Min* mutation, when crossed onto resistant genetic backgrounds, survive long enough to develop tumours other than those of the intestinal epithelium, including mammary carcinomas in 10% of female mice (Moser *et al.*, 1993). Coupled with the observation that *APC* is lost from human breast cancer tissues at a significant rate the *Min* mouse may also prove to be a model for some types of breast cancer (Thompson *et al.*, 1993).

### 1.4.2 Embryonic lethality of the homozygote

Double (homozygous) knock-out animals generated using traditional genetic engineering technologies generally prove unsuitable for the study of tumour suppressor genes which are both vital to embryonic development and involved in tumour development in the adult. In the mouse it was soon revealed that *Apc* was one of these genes, perhaps not surprising due to the central importance of the Wnt pathway in development (Moser *et al.*, 1995). Mice bearing one functioning copy of the murine equivalent of the human *APC* gene could develop to term and were live born, although they developed a large number of polyps of the intestine later in life (Su *et al.*, 1992). Mice homozygous for the *Min* mutation are embryonic lethals; these animals fail to develop past the stage of gastrulation and are reabsorbed well prior to birth, before day 10 of gestation (Moser *et al.*, 1995). This fact, combined with the localisation of *Apc* at the leading edge of migrating cells seems to indicate that the early lethality of these mutants may be due to inability to initiate or to control migration, either through E-cadherin dependant adhesion or through  $\beta$ -catenin related cellular proliferation.

### 1.4.3 Comparison of the *Min* model and human disease

While the *Min* mouse has served as a model of FAP there are nevertheless some significant differences between the human FAP phenotype and symptoms displayed by the *Min* mouse. As mentioned previously, the *Min* mouse tends to develop tumours predominantly in the small intestine, while in the human the disease is characterised by polyp and tumour formation in the large intestine or colon (Moser *et al.*, 1990; Bilger *et al.*, 1996). Additionally the mammary tumours developed by a significant proportion of female *Min* mice do not afflict human FAP patients, although *APC* mutations are found in some human mammary carcinoma (Moser *et al.*, 1993; Furuuchi *et al.*, 2000; Jonsson *et al.*, 2000). On the other hand, FAP patients tend to develop extraintestinal complications not observed in the *Min* mouse; these include desmoid and bone tumours, skin cysts and enlarged retinal pigment cells (Soravia *et al.*, 1997).

Some of these differences in phenotype between human FAP patients and *Min* mice may be caused by the location of the relevant *APC* truncation mutations. A number of alternative murine *Apc* truncation models have been generated through targeted insertion of mutations; these include the APC 1638N mutation, which produces no detectable *Apc* protein from the altered allele (Yang *et al.*, 1997). This line displays a phenotype closer to that traditionally associated with FAP; in addition to the expected intestinal lesions these mice develop the desmoid tumours, cutaneous cysts and display retinal epithelial abnormalities mimicking the symptoms observed in FAP patients (Smits *et al.*, 1998; Marcus *et al.*, 1997). It is also possible that the difference in the predominant location of tumour development, small intestine in the mouse and large intestine/colon in the human may be partially influenced by diet; APC 1638N mice fed on a high fat “western” diet displayed more invasive cancers of the colon, like those found in the human disease, than those on control diet (Yang *et al.*, 1998).

### **1.5 Transgenic and knock-out mice as cancer models**

The modelling of a human disease is often facilitated by the identification of a homologous disorder in a model organism. Because of its small size, short generation and its relatively close evolutionary link to humans the mouse is often the model organism of choice. However, while spontaneous or recessive mutations in mouse lines can sometimes be identified which mimic human disorders, the screening of huge numbers of mice would be required to find “natural” models of human disease. Several techniques have been used to generate models in a more efficient manner, including the application of chemical mutagens to increase the natural rate of mutation followed by screening and identification of affected genes. One notable success of this approach is the *Min* mouse, generated by ENU treatment of male mice prior to breeding (Moser *et al.*, 1990). The generation of mouse mutants through chemical mutagenesis techniques has been complemented by the planned alteration of the mouse genome; altering gene homologues already known to be dysregulated in human disorders.



The ability to modify the genomes of model organisms has given much insight into the developmental and disease roles of a large number of target genes. The study of many human diseases has been made considerably easier through the ability to specifically alter the expression of genes in the mammalian model organism. Congenital genetic disorders such as cystic fibrosis and some forms of diabetes have been successfully mimicked in the mouse through planned genomic modifications, most often through introduced deletion or loss of function mutations in single genes (Dorin *et al.*, 1992; Dorin, 1995; Yamamura *et al.*, 1992; Janson *et al.*, 1996). These altered lines can be crossed with other engineered lines to explore the roles of multiple gene loss in disease progressions or can themselves be further modified (Toft *et al.*, 1999). One of the most important uses for these knock-out or transgenically modified mice is in the study of cancer development and progression.

While tumour development often requires the mutation and/or modification of large numbers of genes it has been discovered that a relatively small number of genes are mutated in a large proportion of cancers; these genes fall into the categories of proto-oncogenes or tumour suppressor genes depending whether gain of function or loss of function mutations lead to initiation of disease. These genes make likely candidates for site specific gene knock-out (loss of function) and/or transgenic (gain of function) manipulation of the mouse genome. Several tumour suppressor genes have been successfully knocked-out in the mouse; these include *p53*, *retinoblastoma (rb-1)*, *p16* and *p19ARF*, *BRCA1/2* and *Apc* among others (Donehower *et al.*, 1992; Clarke *et al.*, 1992; Kamijo *et al.*, 1997; Hakem *et al.*, 1996; Yang *et al.*, 1997). All of these mice display phenotypes mimicking those observed in human patients with germline deficiencies in the human homolog genes; increased tumour incidence either globally or in tissues relevant to the specific gene. Additionally these mice have been used in the development of chemotherapeutic or chemopreventative drugs and have provided a wealth of information on the accumulation of further genetic defects which cause tumour progression (Lipkin, 1997; Smits *et al.*, 1997). Transgenics have been used to allow the overexpression of genes in mouse lines, either through the use of strong promoter or through the inclusion of multiple copies of the transgene simultaneously in the genome. Transgenics have also been used to modulate

knock-out phenotypes generated by site-specific alterations of the mouse genome; by partially compensating for disease causing mutations or rescuing potentially lethal homozygotes (Zhou *et al.*, 1994; Shibata *et al.*, 1997).

### **1.5.1 Introduction of exogenous DNA into mouse oocytes; transgenics**

Transgenic mice were first engineered over 25 years ago; the injection of Simian Virus 40 (SV40) viral DNA into the developing mouse blastocyst resulted in the successful integration of the viral sequences into the host genome, where they were subsequently detectable in all tissues (Jaenisch, 1975). Since this time, the technology has progressed to the injection of more significantly altered DNA sequences into the pronuclei of the fertilised mouse egg; the injected DNA will undergo illegitimate recombination, inserting at one or more essentially random points in the genome, often as concatamers of transgene DNA (Palmiter and Brinster, 1986). Mice produced by this method are screened to determine the number of integration sites and the number of transgene copies which have inserted. The uncertain nature of the integration process involved with transgenic modification can prove to be a significant drawback. It has been noted that the location of integration can have a very significant influence, termed positional effect, on the levels of expression achieved from the transgenic construct. Integration may take place inside the coding sequence of a native gene, disrupting transcription and thereby contributing an additional phenotype to the mouse (Palmiter and Brinster, 1986; Fukushige and Sauer, 1992). An additional limitation of this approach is the inability to efficiently delete expression; while some techniques exist by which transgenic manipulation could reduce or counteract native expression of a particular gene, transgenics cannot efficiently generate null alleles. The ability to efficiently delete or replace specific genomic sequences is termed gene targeting.

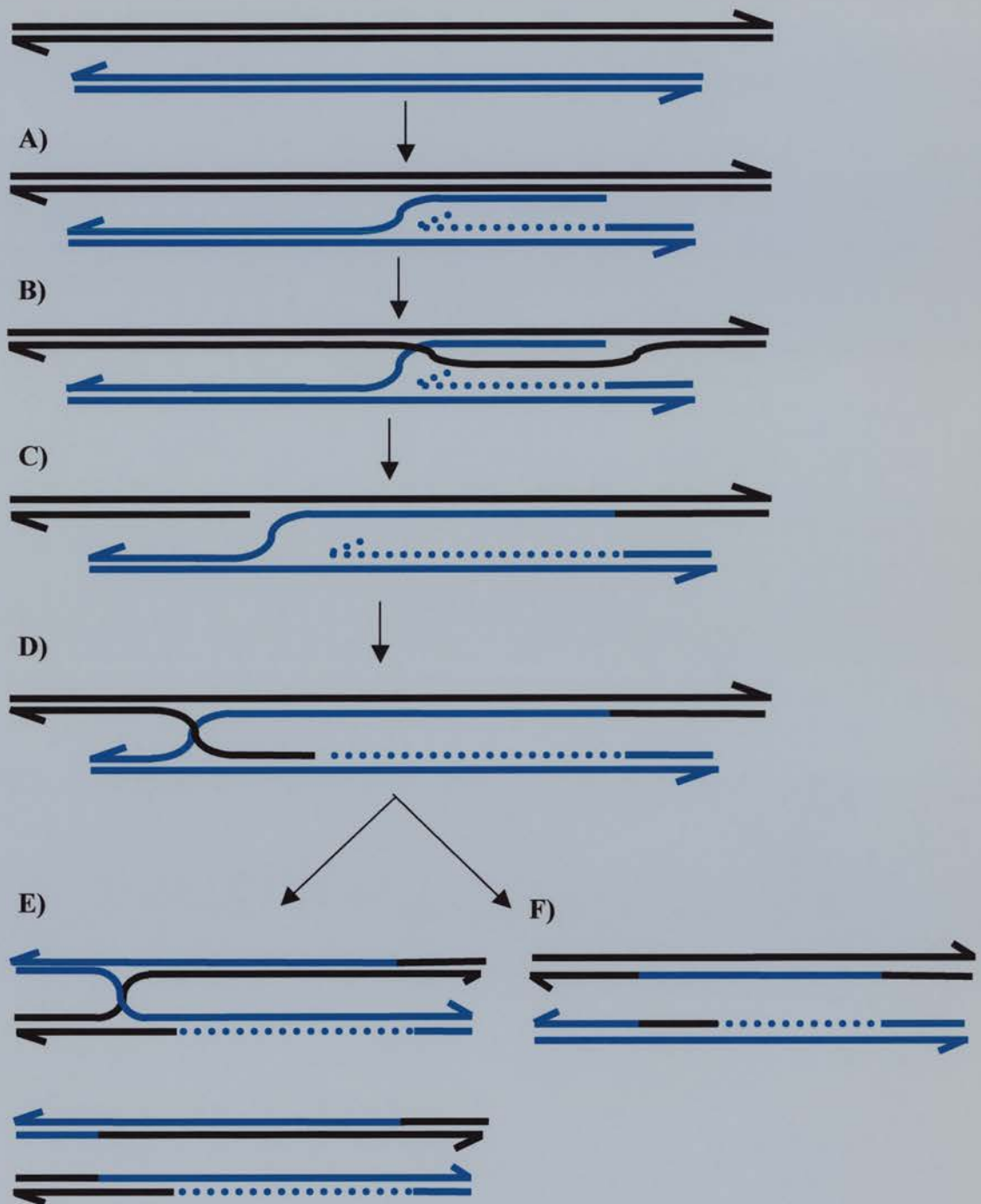


### 1.5.2 Site-specific alteration of the mouse genome to produce disease models; Gene targeting

Altering the genome in a targeted manner requires the harnessing of native cellular mechanisms, some knowledge of the region of genome to be changed and the isolation of cells which can be used to carry the genetic alterations into adult mice. Cells can and do undergo a process of DNA strand exchange between regions of similar sequences referred to as homologous recombination (HR). In eukaryotes this process takes place during gametogenesis where sister chromatids exchange DNA, while in both prokaryotes and eukaryotes HR is involved in the repair of DNA which has suffered mutation or breakage (Bollag *et al.*, 1989). The exact mechanistic processes involved in HR are still not fully understood but a number of models have been proposed. The first model involves the generation of a single stranded DNA molecule with a free end from one of the homologous pair sequences which proceeds to invade the double strand of the target molecule. The displaced DNA strand from the target molecule forms what is referred to as a D-loop; a loop of single stranded DNA which degrades to its own free end which then invades the "attacking" DNA double strand (Meselson and Radding, 1975). The ligation of both of the invading single strands to the free end of the target strands creates a Holliday junction, a structure which can migrate along the DNA molecules, exchanging material and eventually resolve itself in one of two manners leaving behind strand exchange products (see figure 1.5) (Holliday, 1964). The second model of HR requires the double strand breakage of one of the pairs of homologous molecules. One of the free ends thus created invades the unbroken DNA molecule, resulting in the formation of a D-loop. In this case the D-loop formation does not degrade; both strands making up the D-loop act as templates for the cellular DNA replication machinery. The result of this replication is the formation of two Holliday junctions; each capable of migrating and resolving in two ways, leaving a total four possible recombination products (see figure 1.6) (Szostak *et al.*, 1983).

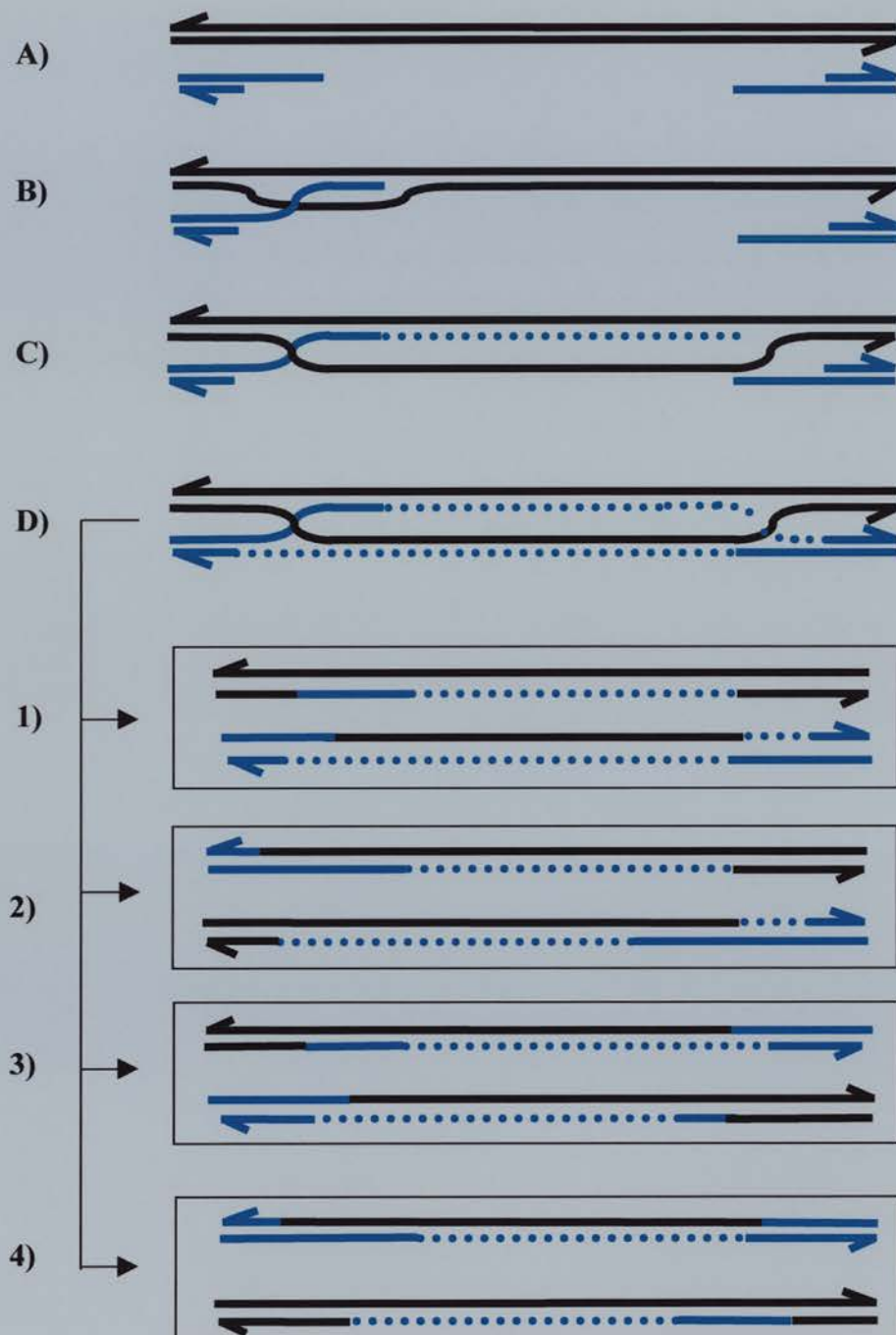


**Figure 1.5: Homologous recombination; the Meselson-Radding model**



**Figure 1.5:** One model of the process of homologous recombination. In this model **A)** recombination is initiated by a single strand nick in one of the participating sequences. The nicked strand is displaced from its complementary sequence by DNA repair synthesis and **B)** invades the homologous double stranded DNA of the other molecule. **C)** The invading strand continues displacing its homologous sequence which is degraded creating an asymmetric heteroduplex. **D)** The free ends of the newly synthesised and partially degraded strands are ligated together forming a Holliday junction which can migrate along the DNA molecules. **E)** The first possible resolution of the Holliday junction produces gene conversion areas as well as crossover segments. **F)** The second mode of resolution of the Holliday junction involves cutting of the crossed strands ( without isomerisation) producing only gene conversion products.

**Figure 1.6: Homologous recombination; the double Strand Break model**



**Figure 1.6:** A second model of the process of homologous recombination. In this model **A)** recombination is initiated by a double strand break in one of the participating sequences. **B)** A single strand from the broken molecule then invades the homologous double stranded DNA of the other molecule. **C)** Repair processes extend the invading single strand, displacing it's homologous sequence from the invaded molecule. **D)** Repair of the broken molecule and ligation of the free ends results in the formation of 2 Holliday junctions. **E)** The two Holliday junctions may resolve in 4 manners involving simple resolution without junction isomerisation (**1**), resolution following isomerisation of one of the junctions (**2,3**) or resolution following isomerisation of both junctions (**4**).

Whatever the exact mechanism of HR, the common factor is the presence in close proximity of two DNA strands of very similar sequence. Because of this requirement for homology the alteration of a particular gene or other genomic structure through HR by necessity requires some knowledge of the sequence of the DNA to be affected. General gene targeting approaches used to generate knock-out mutations of specific genes use arms of homology, regions of sequence identical to the target region, in conjunction with one or more selectable markers and elements designed to interrupt the coding sequence. Successful knock-out HR replaces one or both genomic copies of a target with the selectable marker allowing isolation of successful recombinants. Modifications to this strategy have been developed, some using a second round of HR, to generate subtle mutations in target genes or to remove the selection marker to prevent it contributing a phenotype to the animal ultimately produced (Askew *et al.*, 1993; Stacey *et al.*, 1994).

While the process of HR can generate the desired mutations they are of little use if the cells thus modified cannot generate whole animals. Of primary importance to the creation of models of human disease through specific alterations of target genes was the isolation and development of a culture system for cells which retain totipotency following genetic manipulation. The use of fertilised eggs in the generation of transgenics does not allow for the screening of large numbers of possible recombinants prior to gestation. Embryonal Stem (ES) cells prove to have the properties needed to generate recombinants, screen possible cells and finally generate an entire engineered mouse. These cell lines are isolated from peri-implantation mouse blastocysts; under standard culture conditions these cells differentiate spontaneously into a variety of cell types, underlining their totipotency (Evans and Kaufman, 1981). ES cells plated onto feeder cell layers or grown in media conditioned by Buffalo Rat liver (BRL) cells were able to remain undifferentiated and could be grown, passaged and maintained in culture relatively easily (Smith and Hooper, 1987). Eventually a single cytokine was found to be largely responsible for the maintenance of this state, Leukaemia-inhibiting Factor (LIF); addition of recombinant LIF to ES cell media allows the cells to remain undifferentiated (Williams *et al.*, 1988; Moreau *et al.*, 1988). Importantly, these cells can maintain a normal karyotype and can colonise



various tissues to generate chimeric animals when injected into a mouse blastocyst (Gossler *et al.*, 1986). ES cells are injected or electroporated to introduce the recombination construct and cells having undergone the expected HR are selected and screened. Cells which are found to have the correct genetic alterations are then cultured and introduced into blastocysts; if the introduced ES cells colonise the germline progenitor cells of the chimeric mouse, the genetic alteration can be passed on to generate an animal which contains the manipulated DNA in every cell.

### **1.5.3 Conditional Gene Targeting**

In the search for methods of even finer control of gene expression *in vivo* a number of systems were discovered and developed to allow inducible control of gene knock-out or activation. Two systems have become widely used in a number of model organisms and cell lines; the FLP/frt system, isolated from *Saccharomyces cerevisiae* and the Cre/loxP system discovered in the bacteriophage P1 (Broach and Hicks, 1980; Abremski *et al.*, 1983).

#### **1.5.3.1 Background: The FLP/frt system**

Both the Cre/loxP and the FLP/frt systems rely on specific DNA sequences and on an enzymatic recombinase capable of recognising these sequences. The FLP recombinase, a member of the integrase family, is encoded on the 2 $\mu$ m circle plasmid of *Saccharomyces cerevisiae* and is involved in its replication cycle (Volkert and Broach, 1986a; Volkert *et al.*, 1986b). The specific DNA sequences, frt sites, are 48bp motifs, composed of two 13 base pair direct repeats separated by 1 base, an 8bp spacer element and a third 13 basepair complementary repeat (Meyer-Leon *et al.*, 1984). Three Flp proteins bind to these sequences, in a specific order, binding first to the central 13 basepair motif, followed by binding to the more upstream 13bp repeat and finally to the most downstream repeat (Andrews *et al.*, 1987). This sequential binding introduces a significant bend, greater than 144°, to the DNA (Schwartz and Sadowski, 1990). This bending allows protein interactions between the Flp molecules on paired frt

sites, the end result being a site-specific cleavage and exchange which results in the excision of the DNA sequence between *frt* sites (Amin *et al.*, 1991; Amin *et al.*, 1990; Jayaram, 1994).

### 1.5.3.2 Background: The Cre/LoxP system

The second commonly used recombinase/site system is the Cre/loxP system (Abremski and Hoess, 1984). The Cre recombinase (causes *recombination*) of bacteriophage P1 functions during the replicative cycle of the phage to circularise monomeric genomes out of the concatameric replication product (Segev and Cohen, 1981). The LoxP sites (*locus of crossover*) are composed of two 13bp inverted repeats flanking a central spacer region of 8bp (Hoess *et al.*, 1982; Hamilton and Abremski, 1984). Each of the 13bp repeats binds a Cre molecule during the process of recombination. The spacer region provides a region of DNA bending, allowing two loxP sites on the same DNA molecule to be brought together and made to interact as well as conferring directionality on the loxP site (see figure 1.7a)(Guo *et al.*, 1997; Lee and Saito, 1998). Additionally, the spacer region imparts control over which loxP sites may recombine; several mutant loxP sites have been created bearing altered spacer sequence which will not efficiently react with each other or with wild-type loxP sites but will efficiently recombine with themselves (Araki *et al.*, 1997; Lee and Saito, 1998; Lee and Park, 1998). The functional conformation of the Cre and the loxP sites during recombination reactions consists of 4 Cre molecules (2 per loxP site) and 2 loxP sites brought into parallel positions by bending of the DNA molecule or by alignment of separate DNA molecules (Hoess *et al.*, 1987; Guo *et al.*, 1997). The tetrameric structure then introduces sharp bends in the loxP site spacer regions, thereby positioning the cleavage and ligation sites of the DNA molecule(s) into the active sites of the recombinases (Guo *et al.*, 1997). A set of nucleophilic attacks catalysed by conserved tyrosine residues takes place, ultimately resulting in strand exchange and completed recombination (see figure 1.7b) (Guo *et al.*, 1997). There are a number of possible outcomes of this reaction, depending on the orientation of the loxP sites and the number of molecules involved. When the loxP sites contained in the same DNA molecule are oriented in the same direction the intervening DNA sequence (referred to as

being flanked by *loxP* sites or *floxed*) is circularised and liberated from the parent molecule (see figure 1.8). *LoxP* sites in opposite orientation in the same DNA molecule can result in two reactions; the inversion of the intervening sequence on Cre mediated recombination or the creation of dicentric chromosomes during mitosis (Lewandoski and Martin, 1997). Alternatively, with *loxP* sites on separate DNA molecules, Cre mediated recombination will result in the co-integration of the two molecules. Importantly, this reaction appears to be carried out without the use of ATP or co-factors and functions in a number of organisms, including a wide variety of eukaryotes (Hoess and Abremski, 1985; Sauer, 1987). These characteristics make the Cre/*loxP* system a very useful tool for the manipulation of gene expression and the development of disease models.

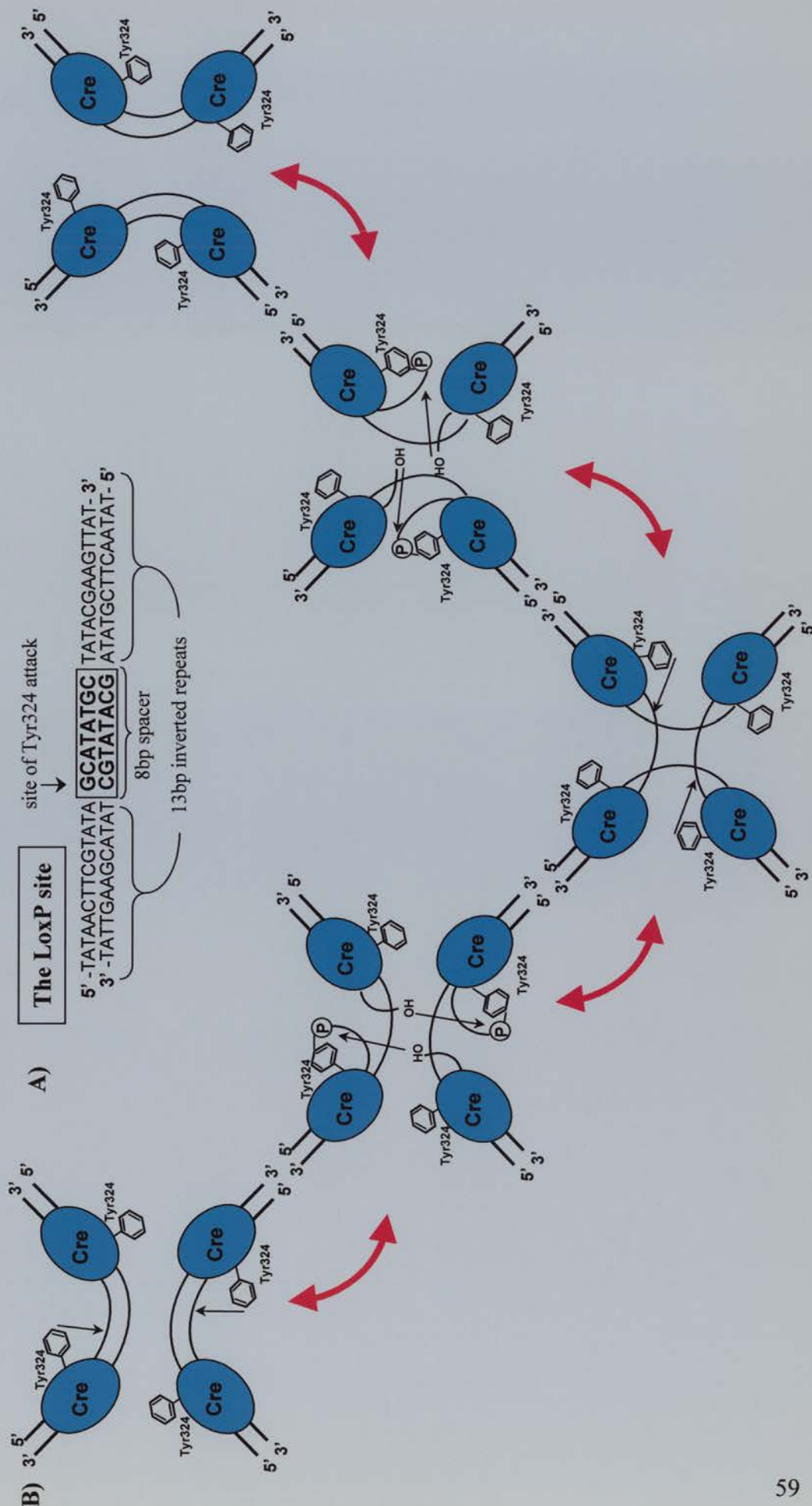
#### **1.5.3.3 Manipulation of the genome using the Cre/LoxP system**

The flexibility of the Cre/*loxP* system makes it useful in a number of different gene modification strategies. The applications with the most potential, perhaps, are those which aim to use the site-specific recombinase to alter the expression of a particular gene in a precise tissue type at a specified time (see figure 1.9). In the drive to achieve this level of control a number of factors have had to be addressed and dealt with; most importantly developing the ability to control the location of Cre activity and/or expression and the generation of strains of model organisms bearing the desired genomic alterations.

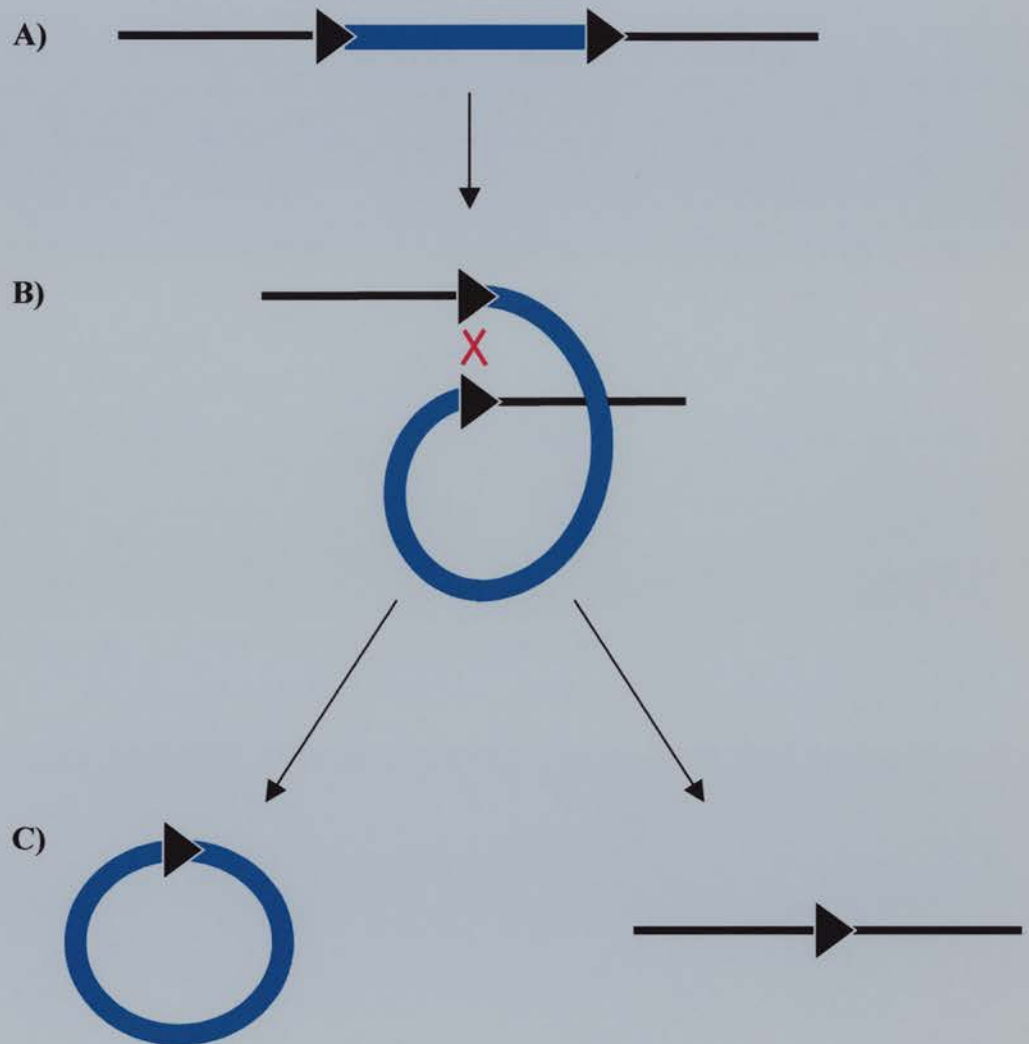
**Figure 1.7:** **A)** The sequence of the native LoxP site. **B)** A representation of the recombination reaction. The tyrosine 324 residue of the Cre recombinase initiates a nucleophilic attack on the phosphate bond of the 5' guanine residue of the loxP site spacer region of one of the DNA strands. The hydroxyl groups of the liberated DNA ends subsequently attack the tyrosine/phosphate bond generating a pair of DNA strand crossovers. The tyrosine 324 residue then attacks the same guanine residue in the remaining DNA strand and the hydroxyl groups again attack the tyrosine/phosphate bond, resolving the DNA strand crossovers.



Figure 1.7: The Cre/loxP recombination mechanics and the loxP sequence



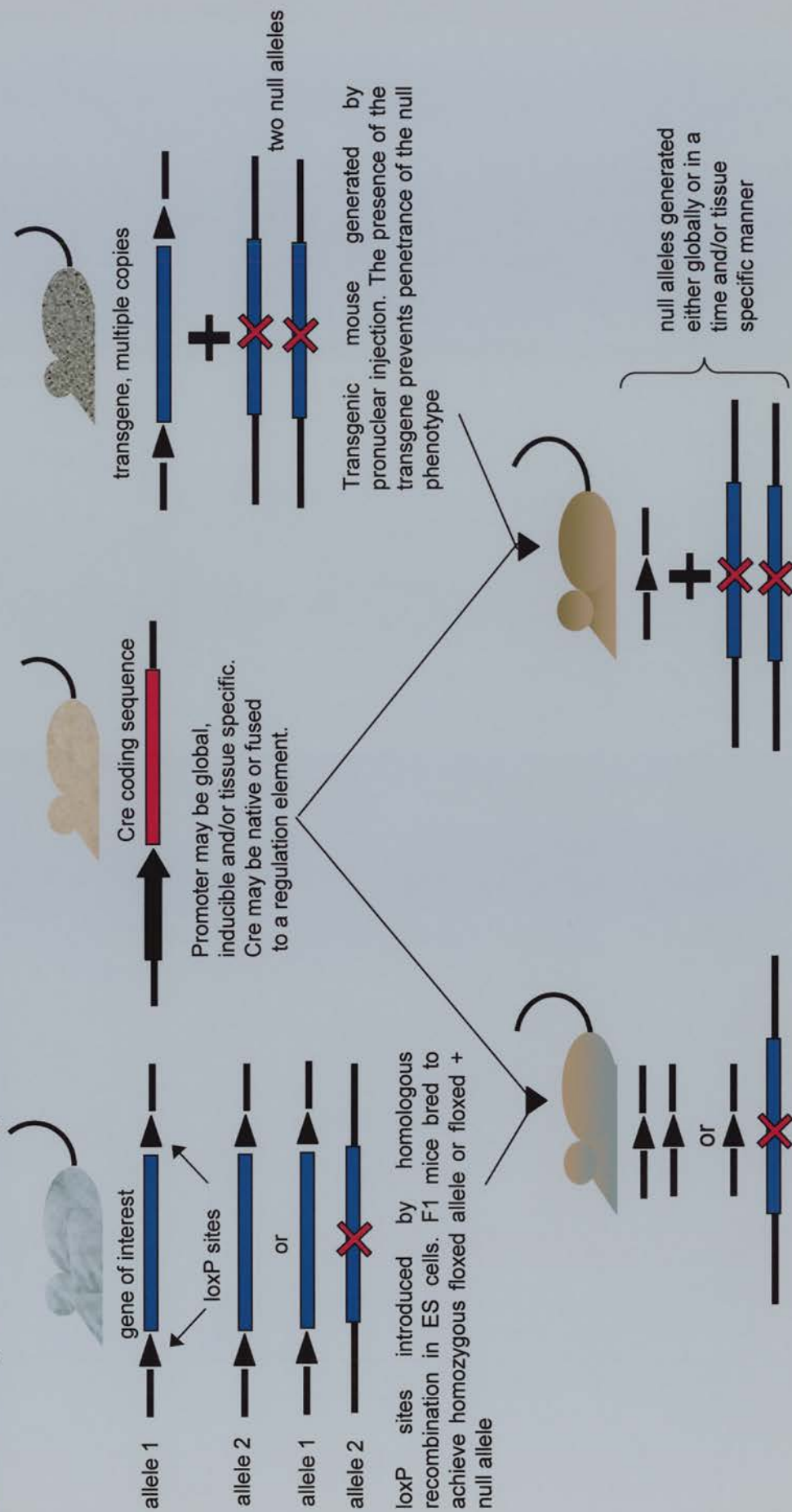
**Figure 1.8: Cre mediated excision of a loxP flanked (floxed) DNA sequence**



**Figure 1.8:** A model of Cre-mediated recombination/excision of a stretch of floxed DNA **A)** The sequence to be excised is flanked by 34bp loxP sites (black arrowheads) in the same orientation within the DNA molecule. **B)** The Cre recombinase brings the loxP sites together in a parallel manner and initiates the recombination reaction (see figure 1.7 for details). **C)** Following recombination the stretch of DNA originally flanked by loxP sites is excised as a closed loop containing one loxP site while the parent molecule retains the second loxP site within its sequence. This process can be reversed leading to Cre induced integration.

**Figure1.9:**The generation of global, tissue specific and/or time specific null phenotypes by Cre recombination.

Figure 1.9: Principle of Cre mediated gene deletion



#### 1.5.3.4 Regulating the expression and/or activity of the Cre recombinase

A number of strategies have been developed to control both the location of Cre expression or delivery as well as its activity *in situ*, many of which have been used *in vivo* to achieve genomic modifications. Some of the tools and techniques employed to achieve this specificity include the delivery of Cre coding sequence in viral systems, the use of tissue-specific or inducible promoters to drive Cre expression and the use of Cre fusion proteins which can be ligand activated.

The delivery of Cre coding sequences directly to tissues *in vivo* or to cell cultures *in vitro* can take place in a number of different manners. The most widely employed method appears to be the use of a viral mediated system. The most successful systems to date involve the use of modified adenovirus vectors. In general a segment of the viral genome is deleted (the E1 or E3 genes most commonly) and replaced with the Cre coding sequence under the control of a strong promoter (Anton and Graham, 1995; Sakai *et al.*, 1995). The resultant virus is replication deficient but still capable of infecting a wide range of cell and tissue types. The infected cells maintain the virus delivered DNA sequences episomally for a short period before being degraded but this transient expression of Cre recombinase may be sufficient to bring about the desired alteration in target gene function. This system has been used with success in CV1 cell lines where high efficiency of infection has been achieved and recombination has been clearly and consistently achieved (Kanegae *et al.*, 1995). *In vivo* the situation is more difficult to control as the presence of an immune system in experimental organisms, as well as a large number of cell types for which adenovirus displays varied tropisms, can result in suboptimal infection efficiencies and alteration of phenotypes (Yang *et al.*, 1994; Yang *et al.*, 1996; Wang *et al.*, 1996). Despite these difficulties a number of mouse lines bearing loxP altered target genes have been successfully infected with these altered adenovirus constructs. The high affinity for hepatocytes displayed by adenovirus makes the liver a prime target for the use of this recombination system. Intravenous administration of Cre encoding adenoviruses has been used to achieve both the activation and deletion of genomic targets in the liver, with a relatively low background of recombination occurring in other tissues (Wakita *et al.*, 1998; Rohlmann *et al.*,



1998; Stec *et al.*, 1999). More specific administration of Cre carrying adenoviruses can also be used: for example, this system has been used to introduce active Cre genes and induce recombination in the epithelium of the murine colorectum through enema application (Shibata *et al.*, 1997). Other virus systems have also been used to deliver Cre genes into target cells; these include Herpes Simplex (HSV) and retrovirus species, each with their own spectrum of efficiency and tissue tropism (Brooks *et al.*, 1997; Choulika *et al.*, 1996).

An alternative to administering Cre coding sequence to genomically altered organisms is the generation of a transgenic line which expresses Cre recombinase in the tissues of interest and crossing this strain to a loxP altered line. The resultant offspring will express Cre, and subsequently undergo genomic recombination reactions, only in the desired locations. A number of mouse lines have been generated to date which express Cre only in particular tissues or sets of tissues, which include thymocytes, cardiac myocytes, pancreatic cells, hepatocytes and the mammary gland among many others (Orban *et al.*, 1992; Agah *et al.*, 1997; Ray *et al.*, 1999; Wagner *et al.*, 1997; Selbert *et al.*, 1998) (see also <http://www.jax.org/pub-cgi/imrpub.sh?objtype=rptquery>). One of the potential benefits of this system, an advantage which many of the Cre/loxP systems have been designed to include, is the ability to knock-out or turn on a gene which may prove lethal if altered globally rather than in an isolated manner. Some difficulties may arise when attempting to use this system, however, as many of the promoters which have specific tissue activity in adult animals may prove to be active in a different location during embryogenesis. This has been noticed in platelet endothelial cell adhesion molecule-1 driven Cre expression and nestin driven Cre expression, where tissue specificity in the adult was accompanied by more global activity in the developing embryo (Terry *et al.*, 1997; Betz *et al.*, 1996). The study of genes involved in the process of embryogenesis, however, may require very early alterations in expression. A number of mouse strains have been engineered to express Cre during development, including Wnt1 promoter driven Cre targeting to the developing neural tube (Danielian *et al.*, 1998).

The use of tissue-specific promoters to drive Cre expression provides a large degree of spatial, and possibly some temporal, control over the recombination process. In order to more specifically regulate the timing of Cre expression or activity in a model organism there are two possible strategies used currently; the use of ligand inducible promoter elements to specifically start or stop Cre expression or the generation of Cre fusion proteins which can be made active through the application of exogenous signals (Sauer, 1998). Inducible promoters exist which will either activate or silence on the presence of an exogenous ligand. A tetracycline dependant Cre expression mouse strain has been developed in which, in the presence of tetracycline, Cre expression is suppressed to very low levels. On withdrawal of the exogenous signal Cre expression is driven in a global manner and genomic recombination takes place (Gossen and Bujard, 1992; St-Onge *et al.*, 1996). Alternatively Cre expression can be driven by the presence of an applied ligand; Cre expression has been driven by promoters induced by the presence of interferon  $\alpha$ , the synthetic hormone RU 486 and the insect hormone ecdysone (Kuhn *et al.*, 1995; Kellendonk *et al.*, 1999; Kellendonk *et al.*, 1996; Feil *et al.*, 1996). In all of these systems the application of the ligand rapidly and reversibly induces the expression of the Cre gene in a global manner.

The alternative ligand activation system involves fusing the Cre coding sequence with an element encoding a ligand-binding domain from an unrelated protein. The fused peptides generally, in the absence of their specific ligand, sterically interfere with the binding of the active sites of the fused Cre protein with the target loxP sites by binding members of the heat shock protein (HSP) family (Picard *et al.*, 1990). In the presence of the proper ligand the HSP is displaced from the fused peptide sequence, allowing the interaction of the Cre active site and the loxP sequences. Several of these fusion proteins have been developed, most often employing native or mutated steroid binding domains (Ligand Binding Domains or LBD's). The constructs include Cre fusions with the native LBD of the glucocorticoid receptor which responds to a number of natural and synthetic steroids as well as an altered glucocorticoid receptor LBD which responds only to the synthetic steroid RU 486, making *in vivo* control more

accurate as natural steroids will not prematurely activate Cre activity (Dias *et al.*, 1998; Kellendonk *et al.*, 1996; Indra *et al.*, 1999). Cre has also been fused to the LBD of the mouse and human oestrogen receptors (CreER), in a mutated form which binds only to derivatives of the synthetic hormone tamoxifen (4-hydroxy tamoxifen most commonly) and is used successfully to generate recombination in double transgenic mice (loxP and CreERTamoxifen) by application of tamoxifen to the skin (Zhang *et al.*, 1996; Feil *et al.*, 1997; Brocard *et al.*, 1997; Vasioukhin *et al.*, 1999). Importantly, the tissue-specific promoter strategy for generation of spatially controlled Cre expression has been successfully combined with the use of inducible Cre fusion proteins to generate a recombination competent mouse strain which can be activated in both a location and a time specific manner (Kellendonk *et al.*, 1999).

#### **1.5.3.5 Site-specific insertion**

As previously mentioned, Cre recombinase will catalyse the integration of two DNA molecules if they each contain one LoxP site. This capability of the recombinase can be used to site specifically insert DNA sequence into, for example, a chromosome which has previously been engineered to contain a loxP site (Sauer and Henderson, 1990). The genomic loxP site(s) can be introduced into an organism through homologous recombination, as mentioned previously, to generate altered sites in known positions or by random integration through illegitimate recombination to generate a number of sites throughout the genome (Sauer, 1998). The main benefit of this technology is the ability to efficiently and reproducibly insert construct DNA sequence into a stable, predetermined genomic location, circumventing or controlling for positional effects. Indeed, this method was used to explore the effects of integration position on transcription from Lactose Z (LacZ) and Neomycin (Neo) resistance cassettes in the mouse genome (Fukushige and Sauer, 1992).

Additionally, avoidance of position effects allowed by site specific insertion of DNA into the genome can be used to more accurately investigate the effects of *cis*-acting elements on promoter activity. The insertion of marker genes and/or enhancer/repressor elements in defined locations allowed detailed study of the



Locus Control Region enhancer element of the human  $\beta$ -globin gene as well as helping to clarify environmental responses of the minimal *p53* promoter (Bouhassira *et al.*, 1997; Bethke and Sauer, 1997).

#### **1.5.3.6 Site specific excision of LoxP flanked DNA sequences**

More common than the use of Cre technology in the targeted insertion of DNA sequences is the use of recombination to excise loxP flanked DNA, either exogenous sequence or endogenous genetic material. This specific deletion of sequence can be used in at least two ways; first to activate transcription from a desired locus and secondly to stop transcription from a gene of interest by physically removing some or all of the coding sequence from the targeted genome.

##### **1.5.3.6.1 Activation of expression through excision (Floxmed stop)**

The strategy for activation of target genes through expression of Cre recombinase generally involves the insertion into the genome of a transgene bearing the gene of interest and a suitable promoter separated by a loxP flanked inactivation cassette (floxmed STOP cassette). This technique was used to target the expression of the SV40 large T antigen to the mouse eye lens; the transgene construct consisting of the lens specific murine  $\alpha$ -crystallin promoter, a floxed cassette consisting of stuffer DNA and a polyadenylation signal and finally the SV40 large T antigen coding sequence (Lakso *et al.*, 1992). The transgenic line generated using this construct did not exhibit SV40 large T antigen expression in any tissue. This line was subsequently crossed to another transgenic line, one expressing the Cre recombinase from a global promoter. The resultant offspring underwent excision of the floxed STOP cassette in all tissues but the expression of the SV40 large T antigen was confined to the lens because of the tissue specificity of the promoter used to drive its expression (Lakso *et al.*, 1992). This Cre-mediated activation of expression has since been used to instigate expression from a number of other genes including hepatitis B genes and *bcl-2* (Shintani *et al.*, 1999; Sato *et al.*, 1998).

#### 1.5.3.6.2 Inhibition of expression through excision (Gene deletion)

Most frequently, the inhibition of expression from an allele using the Cre/loxP technology involves the excision of the floxed gene of interest from the genome. An organism containing a floxed copy of the gene of interest is generated using transgenic or knock-out technology; this strain can then be crossed onto a line bearing a deletion in the gene of interest to generate a model system with one or more functional floxed copies of the gene to be selectively deleted, in addition to either one or two null alleles of the same gene. This line can then be crossed to Cre expressing mice, possibly bearing tissue-specific or ligand activated Cre's to generate what are effectively null alleles in a highly controlled manner.

This technology has been used in the generation of inducible knock-out phenotypes in very specific tissue or cell types. One of the earliest applications of this technique was the ablation of the DNA polymerase  $\beta$  gene in T-cells in the mouse, followed closely by the deletion of the same gene in a ligand inducible manner (Gu *et al.*, 1994; Kuhn *et al.*, 1995). While these mice displayed no observable phenotype, they did effectively demonstrate the feasibility of the applied technology.

Importantly, the degree of tissue and time control afforded through the use of Cre/loxP allows the study of inactivation of genes which would prove lethal in a traditional knock-out model. Included among the genes whose functions have been studied using this system is the murine *Apc* gene. Mouse lines containing an *Apc* gene bearing a loxP flanked exon 14 were generated and backcrossed to generate a line bearing two floxed copies of the *Apc* exon 14 (Shibata *et al.*, 1997). While mice homozygous for truncation mutations of the *Apc* gene are embryonic lethal, failing to develop beyond gastrulation, the floxed exon 14 *Apc* line proved viable and apparently free of phenotype. Upon administration through enema into the colon of adenoviruses encoding Cre recombinase, the floxed sequences were excised, resulting in truncated coding sequences; 80% of these mice rapidly developed colorectal adenomas underlining the tumour

suppressor function of murine *Apc* (Shibata *et al.*, 1997). This mechanism allowed the study of early events in the development of colorectal carcinogenesis and successfully overcame the problem of embryonic lethality posed by homozygosity for traditional mutations in the *Apc* gene.

## 1.6 Project Aims

The use of adenovirus to induce Cre-mediated gene excision gives impressive control over the time and location of gene expression. Even more precise control over this process may be achieved through the development of inducible, tissue-specific Cre expression transgenic mouse lines. One of the goals of this project was to work towards the development of an intestinal specific Cre expressing transgenic mouse. Expression of the *Cdx-1* murine homeobox gene is limited, in the adult mouse, to the stem cells and proliferative compartment of the crypts of the distal small intestine and the large intestine (see section 3.1). Cre expression driven by the *Cdx-1* promoter would confine recombination-mediated excision of floxed sequences to the actively dividing cells of the intestinal crypt. In an effort to produce *Cdx-1*/Cre expression mice, novel *Cdx-1* promoter sequence was to be isolated and characterised *in vitro*, while work was to be carried out towards the development of a bacterial artificial chromosome (BAC) based transgene vector.

One of the primary targets for gene excision using a *Cdx-1* promoter driven Cre recombinase would be the *Apc* gene. A previously generated mouse strain bearing a floxed human APC cDNA expression vector was crossed onto a mouse line bearing the well characterised germline truncation mutation (*Min*) in *Apc*. The resulting progeny were to be characterised by comparison of survival times, tumour spectrum and screening for altered embryogenesis to determine if the introduced transgenic material had a significant influence on the *Min* phenotype.

Additional factors contributing to the development of colorectal tumours in humans include deficiencies in DNA mismatch repair processes. As mentioned previously, recent work has shown that levels of *Apc* are significantly increased in a number of cell types on exposure to the methylating agent MNNG and that this upregulation was dependant on the presence of functional p53 tumour

suppressor protein (Jaiswal and Narayan, 1998). It has been shown that p53 stabilisation following MNNG exposure is, in certain systems, dependent on the presence of the DNA mismatch sensor protein MSH2 (Hickman and Samson, 1999). A further aim of this project was to assess changes following MNNG exposure in the levels of the Apc target protein  $\beta$ -catenin, an indirect measure of cellular Apc levels, in MEFs bearing null mutations in p53 and/or *Msh2*.

# **Chapter 2**

## **Materials and Methods**

## **Chapter 2 – Materials and methods**

### **2.1 Basic cloning Methodology**

#### **2.1.1 Agarose Gel Electrophoresis of DNA**

Agarose gel concentrations used varied from 0.5% to 4% w/v in TBE (see appendix 1) depending on the size of the DNA molecules to be resolved. TBE gels were prepared by boiling an appropriate mass of agarose powder (Sigma) in TBE buffer until dissolved and supplementing with 0.1mg/ml ethidium bromide (EtBr). DNA samples (volumes ranged from 2-50 $\mu$ l) were mixed with one sixth volume of 6x loading buffer (see appendix 1) and loaded directly onto the gel. Molecular weight markers (Markers II, II VI and/or X, Roche) were diluted in loading buffer and run parallel to DNA samples. Electrophoresis was carried out in 1X TBE at 50-150V for periods of 30 minutes to 12 hours and visualised using UV light on a transilluminator (Herolab, gel documentation equipment).

#### **2.1.2 Restriction Digest Analysis of Plasmid DNA**

Suitable amounts of DNA (0.1 $\mu$ g-10 $\mu$ g) were mixed with 0.1 volumes of 10x reaction buffer (supplied by the manufacturer, Roche or New England Biolabs) and 0.1 volumes of 10X Bovine Serum Albumin (BSA) as suggested in the manufacturer's instructions. 0.05-0.1 volumes of appropriate restriction endonuclease (Roche, New England Biolabs), depending on the concentration supplied, were added to the reaction mix and any remaining volume was made up with ultrapure deionised water (dH<sub>2</sub>O) which was prepared using the MilliQ system (resistivity of >18M $\Omega$ -cm); volumes of restriction reactions varied from 10-100 $\mu$ l. Restriction enzymes were stored at -20°C in buffers containing 50% glycerol and kept on ice while in use. Sufficient enzyme was used to ensure complete digestion. 1 unit of enzyme being the amount required to digest 1 $\mu$ g of pBR322 plasmid DNA to completion in 1 hour. The amount of enzyme was kept below 10% of the final reaction volume as glycerol can interfere with enzymatic activity. The digests were incubated for 1-24hrs at the temperature recommended by the manufacturer, varying from 25-50°C (Roche, New England Biolabs). Double digests were either performed in a buffer/temperature combination giving

efficient digestion with both enzymes or were done sequentially with an intermediate phenol:chloroform clean-up step. For phenol:chloroform clean up an equal volume of liquid phenol pre-equilibrated with Tris buffer (Fisher) was added to the reaction mixture, gently vortexed and spun briefly at 10,000g. The upper phase was removed to a clean tube and an equal volume of chloroform-isoamyl alcohol (24:1v/v) was added, gently vortexed and spun briefly at 10,000g. The upper phase was again removed to a clean tube, 0.1 volumes of 3M sodium acetate (pH 5.5) added, followed by 2 volumes of 100% ethanol. The mixture was incubated at -70°C for 30 minutes then spun at 10,000g at 4°C for 10 minutes. The supernatant was poured off and the tube air dried for 30 minutes. A volume of dH<sub>2</sub>O appropriate for the next digestion step was then added and the tube vortexed gently to solubilise the DNA. Digestion reactions were visualised on agarose gels to verify complete cutting.

### **2.1.3 Extraction of DNA Fragments from Agarose Gels**

The QIAEX H Gel Extraction Kit (Qiagen) was used for the purification of DNA fragments from agarose gels. DNA was digested with appropriate restriction endonucleases and electrophoresed on an agarose-TBE gel containing EtBr. The bands were visualised with long-wavelength UV light and excised in a minimal volume from the gel with a clean scalpel and placed in a pre weighed 1.5ml Eppendorf centrifuge tube. The gel slice was weighed and 300µl of supplied QX1 buffer was added per 100mg of gel slice to solubilise the agarose, dissociate DNA binding proteins from the DNA fragment and give a suitable pH to allow absorption of the DNA fragment to the QIAEX II silica particles. 10-30µl of QIAEX II particles, depending on the amount of DNA in the excised gel fragment, were added and the sample incubated at 50°C for 15 minutes with mixing at regular intervals. Following solubilisation and adsorption of the DNA to the QIAEX II silica particles, the sample was centrifuged at 10,000g for 1 minute. The resin pellet was washed once with the solubilisation buffer and twice with the supplied PE wash buffer before being air-dried for 15 minutes. To elute the DNA, the pellet was resuspended in 20µl of 20mM Tris at pH 8.0.



DNA samples extracted from gels in this manner were electrophoresed to determine concentration and used immediately or stored at -20°C.

#### **2.1.4 Dephosphorylation of DNA fragments**

The 5' terminal nucleotides of DNA fragments were dephosphorylated using shrimp alkaline phosphatase, this step is included to increase ligation efficiency by preventing recircularisation of the vector (SAP, United States Biochemical). 1 unit of SAP was added to the DNA in 1x reaction buffer (20mM Tris-HCl pH 8.0, 10mM MgCl<sub>2</sub>) or included in a restriction digest reaction. The reaction was incubated at 37°C for 1 hour then heated to 65°C for 30 minutes to inactivate the enzyme.

#### **2.1.5 Blunting of restriction generated overhangs**

In some cases the 5' or 3' overhangs created in the course of restriction digestion were "blunted" prior to the ligation step. Two protocols were used, depending on whether the overhang to be eliminated was 3' or 5'. In the case of a 5' overhang the large fragment of DNA polymerase I termed the Klenow enzyme (New England Biolabs), was used. 1-5 units of enzyme per µg of DNA and 0.1 volumes of 0.5mM dNTP (Gibco) mix were added directly to a completed restriction digest and incubated at 37°C for 30 minutes. In the case of 3' overhangs the T4 DNA polymerase was used to blunt DNA fragments. This polymerase displays a 3' to 5' exonuclease activity and blunts DNA fragments by removing the 3' bases rather than filling in the overhang. 3-5 units of T4 DNA polymerase per µg of DNA, 0.1 volumes of 4.0mM dNTP mix, 0.1 volumes of 0.5mg/ml BSA and 0.1 volumes of T4 DNA polymerase reaction buffer (see appendix 1) were added directly to completed digestions and incubated for 20 minutes at 12°C followed by heat inactivation by incubation at 75°C for >15 minutes. A fraction of each blunting reaction was gel electrophoresed to check for degradation.

### 2.1.6 DNA Ligation

Concentrations of both vector and insert were calculated by determining the mass of DNA in a given sample volume. A given volume of sample was gel electrophoresed (see above) and compared against suitable mass markers (Low Mass and High Mass Markers, Life Technologies) or by diffusion into EtBr plates. EtBr plates consist of petri dishes containing 1% agarose in 1x TBE buffer supplemented with 0.1mg/ml of EtBr. 1-2µl of sample and 1-2µl of standard DNA diluted to a number of known concentration were pipetted onto plate pre-warmed to 37°C to speed DNA uptake. When the DNA solutions were fully absorbed, the plates were viewed under UV light and comparison between unknown and known DNA concentrations used to determine sample concentration.

Ligations were prepared with 100ng of linearised, dephosphorylated vector and insert was added to a 1:1 or 1:3 molar ratio. Ligation reactions were carried out using two different commercial systems. In the first system the total volume of the ligation reaction was made up to 20 µl with dH<sub>2</sub>O. This was then added directly into a tube of dehydrated Ready-To-Go™ T4 DNA ligase (Pharmacia Biotech). The ligation reaction was incubated at room temperature for 5 minutes, mixed by gentle pipetting and then transferred to a 16°C water bath for overnight incubation. In the second ligation system 1 Weiss unit (equivalent to 67 cohesive end ligation units) of T4 DNA ligase (NEB) and 2µl of 10x reaction buffer (500mM Tris-HCL (pH 7.6), 100mM MgCl<sub>2</sub>, 10mM ATP, 10mM DTT, 50% polyethylene glycol-8000) were added and the reaction made up to a total volume of 20µl with sterile deionised water. Again, ligation reactions were incubated for 5 minutes at room temperature and then overnight at 16°C. Ligation controls were also set up that contained no insert but only dephosphorylated vector.

### 2.1.7 Transformation of Bacteria with Plasmid DNA

Aliquots of TOP10 One Shot™ ultracompetent *E.coli* K12 derived cells (Invitrogen) were removed from storage at  $-70^{\circ}\text{C}$  and thawed on ice. Plasmid DNA was added directly onto the cells and then mixed by pipeting. The cells and the DNA were incubated on ice for 30 minutes, heat shocked in a water bath at  $42^{\circ}\text{C}$  for 45 seconds and returned to ice for a further 5 minutes. 250-500 $\mu\text{l}$  SOC media (see appendix 1) was added to the tube, which was then incubated at  $37^{\circ}\text{C}$  for 1 hour in an orbital incubator at 225rpm.

Electrocompetent cells were generated using an osmotic shock method. Single colonies were picked and grown in an orbital incubator overnight at  $37^{\circ}\text{C}$  in 5ml of L-broth supplemented with the appropriate antibiotic. This started culture was used to inoculate 250-500ml of antibiotic supplemented L-broth which was placed in a  $37^{\circ}\text{C}$  orbital incubator and monitored for Optical Density at 600nm ( $\text{OD}^{600}$ ). Preparations of cells containing the pBAD-Red $\alpha\beta\gamma$  plasmid were supplemented with 0.5% w/v arabinose at an  $\text{OD}^{600}$  of 0.2. When culture  $\text{OD}^{600}$  reached 0.4-0.5, indicative of log growth phase, the culture vessel was placed in an ice-water bath in the cold room and left to cool for 20 minutes. From this point all manipulations were preformed in the cold room with precooled glass and plasticware. Cells were harvested in precooled bottles and rotors by centrifugation at 10,000g for 10 minutes at  $4^{\circ}\text{C}$ . The broth supernatant was removed and discarded and the cell pellets were resuspended in an equal volume of ice-cold  $\text{dH}_2\text{O}$ , followed by recentrifugation at the previous conditions. The cells were then resuspended in 0.5 volumes (relative to the original culture volume) of ice-cold  $\text{dH}_2\text{O}$  and pelleted as before. This pellet was then resuspended in 0.5 volumes of ice-cold 10% glycerol in  $\text{dH}_2\text{O}$  and recentrifuged. This step was repeated in some preparations. The supernatant was removed and the cells resuspended in 0.25 volumes of ice-cold 10% glycerol in  $\text{dH}_2\text{O}$  and again centrifuged. The supernatant was poured off and the cells resuspended in the small volume of remaining 10% glycerol. This resuspension was aliquoted into 50 $\mu\text{l}$  volumes in precooled sterile Eppendorf tubes and snap frozen using dry ice/ethanol baths or liquid nitrogen.

Transformation by electroporation was carried out using a GenePulser (BioRad) connected to a Pulse Controller (BioRad). Aliquots of cells (50µl) were thawed on ice, mixed with 2-5µl of DNA solution. The cell/DNA mixture was then pipetted into prechilled 2mm electroporation cuvettes (Equibio). Electroporations were carried out at 2.5kV, 25µFD and 200Ω. 500µl of SOC media (see appendix 1) was added to the cell/DNA mixture immediately following electroporation. The cell suspension was then incubated at 37°C for 1 hour in an orbital incubator at 225rpm. The transformed bacteria from both protocols were then plated onto Luria Broth (LB)-agar plates (Luria broth supplemented with 2% w/v bacto-agar) containing appropriate selection antibiotics (see appendix 1).

## **2.1.8 Preparation of Plasmid DNA**

### **2.1.8.1 Small Scale Plasmid Preparation**

Small scale preparations of plasmid DNA were purified from transformed bacterial cells with the QIAGEN QIAprep Spin Miniprep kits in accordance with the manufacturer's instructions. This "miniprep" method is a modification of the procedures described in Sambrook *et al.* (Sambrook *et al.*, 1989). A single bacterial clone was picked from a LB-agar plate using a sterile p200 pipette tip. The clone was transferred to a Falcon 2059 tube containing 5 ml of LB (see appendix 1) supplemented with appropriate antibiotics and incubated in an orbital incubator (225rpm) at 37°C overnight. 1.5ml of the overnight bacterial culture was placed in a 1.5ml Eppendorf microcentrifuge tube and spun at 13,000 rpm for 2 minutes. The supernatant was discarded and bacterial pellets were resuspended in 200µl of resuspension buffer (50mM Tris-HCl, pH7.5, 10mM EDTA, 100µg/ml RNase A). RNase A was added to this solution to avoid RNA contamination in later nucleic acid purification steps. Bacterial lysis was accomplished by adding 200µl of lysis solution (0.2M NaOH, 35mM SDS) and mixing by inversion. The lysis reaction is subsequently neutralised by addition of 280µl of neutralisation buffer (1.32M potassium acetate, pH 4.8) and again mixing by inversion. This latter step produces a protein precipitate, subsequently removed by centrifugation at 10,000g for 1 minute. Supernatants were transferred to QIASpin filter tubes and placed in collection tubes; the liquid was

passed through the filter by centrifugation at 10,000g for 1 minute. 500µl of wash buffer (80mM potassium acetate, 8.3mM Tris-HCl pH 7.5, 40µM EDTA, 55% (v/v) ethanol) was added to the QIASpin filter tube followed by centrifugation for 1 minute at 10,000g. The collection tubes were emptied and the filter tubes spun again for 1 minute at 10,000g. QIASpin filter tubes were placed on fresh 1.5ml Eppendorf tubes and 50µl of elution buffer (10mM Tris pH 8.5) was added to each tube. The tubes were incubated at room temperature for 5 minutes. The QIASpin/Eppendorfs were spun at 10,000g for 1 minute. The elute containing the DNA was used immediately or stored at -20°C.

#### **2.1.8.2 Large Scale Plasmid Preparation**

Large-scale isolation of plasmid DNA was carried out using the QIAGEN Plasmid Maxi Kit (Qiagen), the HYBAID Quickflow Maxi Kit (Hybaid) or the Nucleobond Maxi Kit (Clontech) according to the manufacturer's instructions. The general protocol consists of inoculating a conical flask containing 100-500ml of LB, supplemented with the appropriate antibiotics with bacteria from suitable miniprep cultures. The bacteria were grown at 37°C for 12- 16 hours in an orbital incubator at 225 rpm, harvested by centrifugation (10,000g for 10 minutes at 4°C, Sorvall GSA rotor, Centrikon H-401B), resuspended and lysed and neutralised in manufacturer supplied solutions as for the miniprep method. Protein precipitate was removed by centrifugation at 12,000g for 30 minutes at 4°C (Sorvall SS-34 rotor, Centrikon H-401B). Supernatants were added directly to QIAGEN-tips, HYBAID columns or Nucleobond columns and allowed to pass through the DNA binding resin by gravity flow. Columns with bound DNA were washed with the supplied buffers twice prior to elution of DNA. DNA was removed from the column using a supplied elution buffer and 0.7 volumes of isopropanol were added. The DNA was collected by centrifugation at 12,000g for 15 minutes at 4°C (Sorvall SS-34 rotor, Centrikon H-401B). DNA pellets were washed with 70% ethanol and recentrifuged. The final DNA pellet was dissolved in 100-250µl of TE buffer (10mM Tris-HCl, pH7.5, 1mM EDTA) or 20mM Tris pH 8.0, in accordance with manufacturer's instructions.

### 2.1.8.3 Preparation of BAC DNA

Preparations of BAC DNA follow a protocol developed by *Davidson et al.* (Davidson *et al.*, 1999). BAC isolation is carried out using the Q-tip 500 (Qiagen) plasmid extraction kit. Each extraction requires eight Q-tips and an additional 40ml each of Resuspension, Lysis and Neutralisation solutions (see appendix 1) and 180ml extra buffer QC (see appendix 1). Each extraction requires 2.5 litres of bacterial culture, which should be grown from a single colony derived from a plate that has been freshly streaked out from the original bacterial stock. A single colony was picked and used to inoculate 5ml of L-broth containing 20µg/ml chloramphenicol. This starter culture was grown at 30°C for 6 to 8 hours and subsequently used to inoculate 500ml of L-broth again supplemented with 20µg/ml chloramphenicol and allowed to grow at 30°C over night. Cells were centrifuged, the supernatant removed and, if required, stored at -20°C.

Cells were resuspended in a total of 120ml of Resuspension buffer, lysed by the addition of 120ml of Lysis buffer and incubated at room temperature for 5 to 10 minutes with gentle mixing by inversion. The lysis reactions were stopped by the addition of 120ml of Neutralisation buffer, followed by inversion mixing. The resulting solution was divided into centrifuge bottles and spun at 7,000g for 30 minutes in a Sorvall RC-5B rotor at room temperature to precipitate proteins.

During this centrifugation step 8 Q-tip columns were pre-equilibrated with 10ml each of buffer QBT (see appendix 1). The supernatant from the centrifugation was filtered through four layers of gauze and 45ml loaded onto each column. The columns were washed with three aliquots of 30ml of buffer QC (see appendix 1). Columns were eluted with 15ml of buffer QF (see appendix 1) and the eluates combined in 50ml conical tubes. BAC DNA was precipitated by the addition of 20ml of room temperature isopropanol to each tube and centrifuged at 3250 rpm in a Sorvall RT7 with swinging bucket rotor for 30 to 60 minutes at 4°C. The pelleted BAC DNA in each tube was washed with 20ml of 70% ethanol and recentrifuged for 15 minutes. Pellets were drained and air-dried before being



resuspended in 195µl of ultra-pure water. To reduce genomic DNA contamination, 10µl of 10mM ATP solution, 25µl of 10x Plasmidsafe buffer and 5 to 10µl of Plasmidsafe enzyme were added to each 195µl resuspension and allowed to digest overnight at 37°C while the BAC DNA solubilises. The resulting BAC DNA solutions were combined into two Eppendorf tubes and extracted once with an equal volume of phenol/chloroform and ethanol precipitated (see section 2.1.2). The BAC DNA pellets were each dissolved in 200µl TE (see appendix 1) and allowed to dissolve overnight.

#### **2.1.8.4 Preparation of bacterial glycerol stocks**

Single clones were used to inoculate 5ml of Luria Broth supplemented with the appropriate antibiotic and grown overnight at 37°C in an orbital incubator. Bacteria were harvested by centrifugation at 13,000rpm for 3 minutes in a bench-top microfuge followed by removal of excess media. The pellet was resuspended in 750µl of autoclaved 50% Luria Broth:50% autoclaved glycerol and placed into a CryoTube<sup>TM</sup> (Nunc) prior to storage at -70°C.

#### **2.1.9 Sequencing of Plasmid/PCR DNA**

Where only short stretches of sequence needed to be determined, plasmid DNA was sequenced using the T7 Sequenase® quick denature plasmid sequencing kit (Amersham Life Sciences). This protocol is based on the chain termination method originally described by (Sanger *et al.*, 1977). Sequencing involves denaturing 0.5-3µg of plasmid DNA, in a volume of 8µl, by the addition of NaOH (supplied by the manufacturer) to a final concentration of 0.2M. 2pmol of sequencing primer (chosen using the Primer3 web software, see below) was added to the reaction mix. The samples were then incubated for 10 minutes at 37°C before being incubated on ice. The reaction mix was then returned to neutral pH by the addition of HCl (supplied by the manufacturer) to a concentration of 0.2M. Finally 2 µl of 5x plasmid reaction buffer (see appendix 1) was added and the samples were left to anneal for 10 minutes at 37°C and once again placed on ice. To the cooled reaction mixture 0.1M dithiothreitol, 1x



labelling mix (see appendix 1), 5 $\mu$ Ci ( $^{35}$ S) dATP (0.5 $\mu$ l) and 1  $\mu$ l of Sequenase® enzyme (diluted 1:8 in enzyme dilution buffer, see appendix 1) were added. The reaction mixtures were then vortexed gently and incubated at room temperature for 5 minutes. 3.5 $\mu$ l of this labelling reaction was then added to each of 4 0.5ml Eppendorf tubes, each containing 2.5 $\mu$ l of one of the dideoxynucleotide termination mixes (see appendix 1) and incubated at 37°C for 5 minutes. The chain termination reactions were each stopped by addition of 4 $\mu$ l of sequencing stop solution (supplied by manufacturer, see appendix 1). Individual chain termination reactions were then run out side by side at 70W for 1-2 hours on prewarmed 6% acrylamide, 6M urea 0.4mm sequencing gels (see appendix 1). 1xTBE or 1x glycerol tolerant running buffer were used for electrophoresis. Gels were washed twice with 10% methanol/10% acetic acid in dH<sub>2</sub>O and removed to 3MM Whatman filter paper, vacuum dried at 80°C and placed into film cassettes with X-ray film (Kodak BioMaxMR X-ray film, Kodak Eastman Corporation) for 2-3 days.

Where longer stretches of sequence were to be determined automated fluorescence label sequencing was carried out by DNAShef (Royal Infirmary of Edinburgh, Haematology Department).

#### **2.1.10 Southern Blot/Dot Blot hybridisation**

For southern blot analysis ~10  $\mu$ g of the DNA to be probed was digested with the appropriate restriction enzymes (see above). Following digestion, the samples were subjected to electrophoresis on a TBE-agarose gel (see above) at low voltage (40V) overnight at 4°C. Images of the electrophoresed DNA were captured on a UV transilluminator (see above) for later comparison. Transfer of the DNA to the nylon membrane was improved by partial depurination ('acid nick') of the DNA prior to transfer. Partial depurination was achieved by soaking the gel for 10 minutes in 0.2M HCl. The gel was then rinsed in deionised water and the DNA was transferred immediately. DNA was transferred from the gel onto a positively charged nylon membrane (Zeta-Probe® GT, BioRad). 0.4M sodium hydroxide was used as the transfer buffer and transfer

was carried out overnight. After transfer the membrane was subsequently rinsed in 2xSSC (see appendix 1), air-dried UV crosslinked (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corp.). Dot blot membranes for Bacterial Artificial Chromosome (BAC screening) were obtained from Research Genetics. Membranes were stored between sheets of Whatman 3MM paper at room temperature until required.

25-50ng of clean probe DNA in 11µl of dH<sub>2</sub>O (gel extracted restriction fragment or PCR product) was radiolabelled with <sup>32</sup>P-dCTP using High Prime enzyme (Boehringer Mannheim) by addition of 4µl of enzyme and 5µl of <sup>32</sup>P-dCTP followed by incubation at 37°C for 20 minutes. Labelled probe was separated from unincorporated nucleotides using a Sephadex G50 column (Pharmacia). 20µl carrier salmon sperm DNA (10mg/ml) was mixed with radiolabelled probe and added to the column. Unincorporated nucleotides were eluted by addition of 400µl TE (10mM Tris pH 8, 1mM EDTA) buffer which was allowed to enter the column by gravity flow and a further 400µl TE was added to elute the purified probe. The elutes were collected in 1.5ml Eppendorf tubes and the quality of radiolabelling was assessed by Geiger Counter comparison of the counts/minute recorded from both elution products. Labelled probe was only used if it had incorporated >50% of the radionucleotides.

Hybridisation buffer (see appendix 1) was preheated to 68°C to ensure all components were in solution. 10mls buffer was transferred to a hybridisation tube (Hybaid) containing the nylon membrane to allow the membrane to equilibrate. Membrane and buffer were incubated with rotation at 68°C for 1 hour in a hybridisation oven (Hybaid). Following this prehybridisation step, the 400µl of TE containing the appropriate radiolabelled probe was added to the hybridisation tube. The membrane was incubated with the radiolabelled probe in hybridisation buffer overnight at 68°C. The membrane was then washed twice for 15 minutes in 2xSSC supplemented with 0.1% sodium dodecyl sulphate (SDS), once for 30 minutes in 1xSSC/0.1% SDS and once for 10 minutes in 0.1xSSC/0.1% SDS, all at 68°C. Following these washes the membrane was removed from the hybridisation tube, sealed in a plastic bag, and exposed to MR

X-ray film (Kodak) in an autoradiography cassette. Autoradiography cassettes were stored at -70°C for 2-5 days, depending on signal strength. Autoradiography films were developed (Hyperprocessor, Amersham) and compared with the gel image or dot blot grid supplied by the manufacturer.

## 2.2 PCR reactions

### 2.2.1 Reaction Conditions

PCR reactions were carried out in 0.5ml thin-walled microcentrifuge Eppendorf tubes in a standard thermocycler (Hybaid Omnigene, Hybaid). Reaction volumes of 25µl and/or 50µl were used and all reactions were overlaid with 25µl of autoclaved paraffin oil. All reagents were obtained from Life Technologies. Four reaction master mixes of varying magnesium concentrations and additives were used; their recipes are listed in the following table.

**Table 2.1: PCR master mix recipes**

	<b>1.0 mM Mg reaction mix</b>	<b>1.5 mM Mg reaction mix</b>	<b>2.0 mM Mg reaction mix</b>	<b>1.5 mM Mg + DMSO reaction mix</b>
dH <sub>2</sub> O	195.5µl	193µl	190.5µl	180.5µl
10x PCR buffer	25µl	25µl	25µl	25µl
W1 detergent	12.5µl	12.5µl	12.5µl	12.5µl
DMSO	0µl	0µl	0µl	12.5µl
Mg (50 mM)	5µl	7.5µl	10µl	7.5µl
dNTP (10 mM)	8µl	8µl	8µl	8µl
primer (100µM)	2µl each	2µl each	2µl each	2µl each

### **2.2.2 Primer Design**

For most reactions primers were chosen from a number of primer pair options supplied by Primer3, a web base primer design software program (<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>). Primers were chosen to avoid secondary structure and primer dimerisation and to have very similar melting temperatures. Some sets of genotyping primer pairs and reaction conditions were taken from the referenced publications. Primer pairs and the optimal reaction conditions are listed in the following table: Hot start reactions included the addition of 1.5µl of 1:10 dilution of Taq polymerase in dH<sub>2</sub>O once tube temperature had reached 94°C.

**Table 2.2: PCR reaction primers and conditions**

PCR Reaction	Primer pair (5'-3') and product size	Reaction mix	cycle parameters
<i>Cdx-1</i> promoter, A-F reaction	Left: TTACCATTACCAACAGGATC Right: CTGCGCCGCATCCATTCGTAG product size: 1550 bp	1.5mM Mg	94°C -5 minutes 55°C -30 seconds 72°C -1 minute 32 cycles 94°C -30 seconds 72°C -5 minutes
<i>Cdx-1</i> exon1	Left: GTGGGCTATGTGCTGGACAA Right: CATCCATTCGTAGGGCTCC product size: 400bp	1.0mM Mg	Hot start 94°C -5 minutes 59°C -30 seconds 72°C -1 minute 34 cycles 94°C -30 seconds 72°C -5 minutes
<i>Cdx-1</i> promoter, short reaction	Left: GTCTCCTTTTGAACCCCTC Right: GGAGTCCTTGTCAGCACAT product size: 278 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -1 minute 32 cycles 94°C -30 seconds 72°C -5 minutes
Mutant LoxP flanked Neomycin/Kanamycin	Left: ATAACCTTCGTATAGCGTACAT TATACGAAGTTATAGGGAAG AAAGCGAAAGGAG Right: ATAACCTTCGTATAATGTACGC TATACGAAGTTATGGTGGGGA AAAGGAAGAAAC product size: 1802 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -2 minutes 32 cycles 94°C -30 seconds 72°C -5 minutes
pBAD- $\alpha\beta\gamma$ screen	Left: ATGGACGTAAACTACGGC Right: CTACCTGGCTCAAGTCCAG product size: 695 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -1 minute 32 cycles 94°C -30 seconds 72°C -5 minutes
BAC 219 recombination screen, Neomycin/Kanamycin into <i>Cdx-1</i> exon1	Left1: TTCTCCACCTGTAACCCAG Left2: TGCTCCTGCCGAGAAAGTAT Right: CCACTCTCGAGAAGCCAATC product size: 1550bp with recombination 950bp without recombination	1.5 mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -2.0 minutes 32 cycles 94°C -30 seconds 72°C -5 minutes
<i>Cdx-1</i> primer 10-11 reaction	Left: CTACCCAGAATGGCGGTCTA Right: GGTCTGAGCGGAGATACGTC product size: ~1550 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -1.5 minutes 32 cycles 94°C -30 seconds 72°C -5 minutes
<i>Cdx-1</i> primer 12-13 reaction	Left: ATCCCTCTAAAGGGCTGGAG Right: GAAGCAGCCTCTTCCTTTCA product size: ~500 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -1 minute 32 cycles 94°C -30 seconds 72°C -5 minutes

**Table 2.2: PCR reaction primers and conditions (cont.)**

<i>Cdx-1</i> primer 14-15 reaction	Left: TTTTTCCTTGATGGTGGTC Right: TTCTGGGTTCAGTGAGAGC product size: ~1100 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -1.5 minutes 32 cycles 94°C -30 seconds 72°C -5 minutes
<i>Cdx-1</i> primer 14- <i>Cdx-1</i> Down reaction	Left: TTTTTCCTTGATGGTGGTC Right: GCTTGATGATCCTCGTTGGT product size: ~ 1700 bp	2.0mM Mg	94°C -5 minutes 58°C -30 seconds 72°C -1.5 minutes 32 cycles 94°C -30 seconds 72°C -5 minutes
<i>p53</i> genotyping (Clarke <i>et al.</i> , 1993)	Left: GTGGTGGTACCTTATGAGCC Right: CAAAGAGCGTTGGGCATGTG Right 2: CATCGCCTTCTATCGCCTTC product size: 642bp (wt) / 510bp ( <i>p53</i> )	note: master mix uses 100mM dNTP's	94°C -5 minutes 62°C -1 minute 72°C -1 minute 30 cycles 94°C -1 minute 72°C -5 minutes
<i>Msh2</i> genotyping (Toft <i>et al.</i> , 1999)	Left: CGGCCTTGAGCTAAGTCTATT ATAAGG Right 1: GGTGGGATTAGATAATGCCTG CTCT Right 2: CCAAGATGACTGGTCGTACAT AAG product size: 164bp (wt) / 194 bp ( <i>Msh2</i> )	note: master mix uses 100mM dNTP's	94°C -5 minutes 60°C -2 minutes 72°C -2 minutes 30 cycles 94°C -2 minutes 72°C -5 minutes
<i>Min</i> genotyping (Luongo <i>et al.</i> , 1994)	Left: TCTCGTTCTGAGAAAGACAGA AGCT Right: TGATACTTCTCCAAAGCTTG GCTAT product size: 144bp ( <i>Min</i> ) / 123bp (wt)	note: master mix uses 100mM dNTP's	94°C -5 minutes 60°C -2 minutes 72°C -2 minutes 30 cycles 94°C -30 seconds 72°C -5 minutes followed by HindIII digestion
APC transgene screen (Cooper, 2000)	Left: CTCTTCATCGGGAATGC Right: CAGGAACAGCTCTGAC product size: 240 BP	note: master mix uses 100mM dNTP's	94°C -5 minutes 55°C -30 seconds 72°C -30 seconds 34 cycles 94°C -30 seconds 72°C -5 minutes
<i>Mom1</i> allotyping (Santos <i>et al.</i> , 1998)	Left: ACAGGTCCAAGGGAACAT Right: TCTGTGGCATCCTTGGG product size: 118 bp (CBA)/138 bp (B6)	2.0mM Mg	94°C -5 minutes 55°C -30 seconds 72°C -1 minute 32 cycles 94°C -30 seconds 72°C -5 minutes followed by BamHI digestion

## 2.3 Tissue culture

### 2.3.1 Derivation of cell lines

The human colorectal cell lines used were obtained from the European Cell Culture Collection. These include DLD-1, HT 29, HCT116 and SW 480 cell lines (see table 2.3 for references and ATCC numbers). HM-1 ES cells were obtained from Dr. Nathalie Sphyris. Murine embryonic fibroblasts were generated from various mouse lines on site. Pregnant female mice were sacrificed at day 12-14 of gestation and the whole womb removed and placed in a petri dish containing PBS (see appendix 1). The embryos were then individually removed from the womb and separated from the placental tissue using watchmaker forceps, dipped briefly in 70% ethanol to remove contaminating bacteria and placed in a fresh petri dish. Each embryo was mechanically disrupted using new, sterile scalpels, with a small tissue sample taken for genotype analysis where necessary. The disrupted embryo was then suspended in 5ml of 0.2µm filter sterilised 2.5% trypsin (DIFCO) in dH<sub>2</sub>O, prewarmed to 37°C. This suspension was incubated at 37°C for 5 minutes and gently mixed by regular inversion. The trypsin was then neutralised by the addition of 10ml of prewarmed media (see section 2.3.2). The suspension was then centrifuged at 1000g for 3 minutes to pellet the separated cells. The supernatant was carefully removed either by pouring or by pipette and the cell pellet resuspended in 10ml of prewarmed media. This was then poured into a 10cm petri dish and placed in a tissue culture incubator (see section 2.3.2).

Primary wild type murine colonocytes were supplied by Robert Morris and were isolated using the protocol developed by Booth *et al.* (Booth *et al.*, 1995). Briefly, to seed four 24-well plates 3 adult mice were sacrificed and their colons excised, flushed and placed into Hanks' Balanced Salts Solution (HBSS). The colons were slit open longitudinally and dissected into small pieces. The colon fragments were washed in 50ml HBSS and then mechanically disrupted in a petri dish with a scalpel. The disrupted tissue was then placed into 25ml culture media (see table 2.3) supplemented with 75U/ml collagenase and 20µg/ml dispase and incubated with gentle shaking for 2-3 hours at 37°C. The tissue solution was then



transferred to a 30ml Universal tube and allowed to gravity settle for 45 minutes to pellet undigested tissue. The top 22ml of the suspension, containing the isolated crypts was removed and retained. The remaining 3ml of suspension was supplemented with 22ml of fresh media and resuspended and pelleted to collect any remaining crypts. The collected crypt suspensions were diluted with an equal volume of culture media supplemented with 2% D-sorbitol to aid buoyancy and centrifuged at 200-300rpm in a Harrier 18/80 centrifuge (Sanyo) for 2 minutes. The supernatant, containing single cells and bacteria was removed and the pellet, containing whole crypts was resuspended in a further 50ml of D-sorbitol supplemented media and recentrifuged; this washing step was repeated 4-5 times. The pellet was finally resuspended in an appropriate amount of culture media and plated into collagen coated tissue culture plasticware. Collagen coatings were added to the culture vessels by adding sufficient 200µg/ml acidified collagen solution to the vessel to cover the culture surface and leaving it overnight. The collagen solution was aspirated and the vessel allowed to air dry prior to crypt plating. Colonocyte cultures were fed every 3-4 days when half of the media volume was removed and replaced with fresh culture medium.

### **2.3.2 Thawing, media and incubation conditions**

To thaw cell lines for culture, aliquots of cells stored in CryoTube™ vials (Nunc) were thawed with agitation in a prewarmed 37°C waterbath. The rapidly thawed cell suspensions were then dispersed into 10 mls of the appropriate culture media (see table 2.3) and spun at 1000rpm for 4 minutes. The media was aspirated and the cells resuspended in a further 10mls of media. This suspension was placed into the appropriate culture vessel.

All cell lines used were incubated at 37°C in 5% CO<sub>2</sub> and 100% humidity in Sanyo CO<sub>2</sub> incubators. Cells were cultured in plasticware (flasks and petri dishes) supplied by Nalge-Nunc or Greiner. ES cell culture vessels were pretreated by incubation for 30 minutes with an appropriate volume of 0.1% swine skin gelatine (Sigma) in PBS. All cells were removed from plasticware for passaging by removal of culture media, washing with warm PBS and application

of an appropriate amount (enough to just cover the cells) of 0.25% Trypsin/EDTA solution (Life Technologies) prewarmed to 37°C. The cells were incubated in the trypsin solution until the monolayer detached and fresh prewarmed culture media was added to neutralise the trypsin. The resulting cell suspension was diluted and split into the appropriate number of culture vessels. The media used for each of the cell lines used are listed in the following table:

**Table 2.3: Cell culture media**

Cell line	Culture Media	Reference and ACTT number
HCT 116	BHK-21 medium with 2.0mM L-glutamine (Gibco) + 10% foetal calf serum (FCS, Sigma)	ATCC CCL 247 (Brattain <i>et al.</i> , 1981)
SW 480	Leibovitz's medium with 2mM L-glutamine (Gibco) + 10% FCS (Sigma)	ATCC CCL 228 (Leibovitz <i>et al.</i> , 1976)
DLD-1	RPMI 1640 medium with 2mM L-glutamine (Gibco) + 1.5g/L sodium bicarbonate + 4.5g/L glucose + 10mM HEPES + 1.0mM sodium pyruvate (Gibco) + 10% FCS (Sigma)	ATCC CCL 221 (Dexter and Hager, 1980)
HT-29	McCoy's medium with 1.5mM L-glutamine (Gibco) + 10% foetal calf serum (FCS, Sigma)	ATCC HTB-38 (Fogh, 1975)
murine embryonic fibroblasts	BHK-21 medium with 2.0mM L-glutamine (Gibco) + 1x MEM Non-essential amino acids (Gibco) + 1.0mM sodium pyruvate (Gibco) + 10% FCS (Sigma)	N/A
wild type primary mouse colonocytes	DMEM medium (Gibco) + 0.25U/ml Insulin + 100U/ml penicillin + 30µg/ml streptomycin + 25µg/ml gentamycin + 1% batch tested FCS (Sigma)	N/A
HM-1 murine ES cells	BHK-21 medium with 2.0mM L-glutamine (Gibco) + 1x MEM Non-essential amino acids (Gibco) + 1.0mM sodium pyruvate (Gibco) + 5% FCS (Sigma) + 5% normal calf serum (NCS, Gibco) + 0.2mM β-mercaptoethanol (Gibco) + 1:1000 Leukaemia-inhibiting factor (LIF) (Gibco)	N/A (Kitani <i>et al.</i> , 1996)

### **2.3.3 Harvesting cells for protein**

Cells to be harvested for protein analysis were washed twice with ice-cold PBS after removal of culture media. A minimal volume (0.5-2ml) of ice-cold RIPA buffer (see appendix 1) was then added to the cells. The cells were removed from the plastic surface into the RIPA buffer using a cell-scraper (Nalge-Nunc) and the resulting suspension was collected in a 1.5ml Eppendorf tube. The tubes were then immediately frozen, either in an ethanol/dry ice bath or in liquid nitrogen. The suspension was thawed on ice and the cells and DNA mechanically disrupted by being passed through a succession of gradually narrower syringe needles (21,23,25 and 26 gauge). Samples were stored at -70°C.

### **2.3.4 Chemical treatment of cells**

Cells to be treated with MNNG were serum starved (culture media supplemented with 0.5% foetal calf serum) for 18 hours prior to the addition of the methylating agent. 0.050g of MNNG were dissolved in 0.340ml of dimethylsulfoxide (DMSO) to generate a 1.0M solution. This solution was diluted in culture media supplemented with 0.5% serum to concentrations of 25µM and 50µM; untreated cells were fed culture media containing similar amounts of MNNG-free DMSO. Cells were left incubating at 37°C for 15 hours and then washed and harvested as previously outlined. Media containing MNNG and PBS used to rinse MNNG treated cells were collected in disposable plastic universal tubes (Falcon) and sealed with parafilm before being double sealed in collection bags prior to incineration. Plasticware that came into contact with MNNG was double bagged before being disposed of by incineration.

## **2.4 Immunoblotting**

### **2.4.1 Tissue sample preparation**

Intestine samples were removed from the sacrificed animals and frozen in 1.5ml Eppendorf tubes in liquid nitrogen. Samples to be prepared were placed in a mortar prechilled with liquid nitrogen and ground into a fine powder with a

pestle. This powdered tissue was placed in a 25ml universal tube and 2ml of ice-cold RIPA (50mM NaCl, 1% NP-40, 12mM deoxycholate, 3mM SDS, 50mM Tris pH 7.5) buffer was added. The samples were then mechanically homogenised and centrifuged in 1.5ml Eppendorf tubes at 10,000g for 2 minutes at 4°C to pellet cell debris. The supernatant was removed to a fresh Eppendorf tube and passed through a 25-gauge syringe needle to shear remaining DNA. Samples were stored at -70°C.

#### **2.4.2 Protein quantification**

Quantification of protein concentration in cell and tissue lysates was carried out using the BCA protein Assay Kit (Pierce). Briefly, fifty parts BCA reagent A (see appendix 1) were mixed with one part BCA reagent B (see appendix 1) to create the working BCA reagent solution. A protein concentration standard curve was generated by preparing a set of dilutions (5, 2.5, 1.25, 0.625, 0.3125 and 0 mg/ml) of Bovine Serum Albumin (BSA, supplied) in sterile test tubes. 0.1ml of 1:10 dilution of each sample to be assayed was placed in a test tube. 2.0ml of the working BCA reagent solution was added to each test tube, mixed and incubate at 37°C for 30 minutes. At the end of this incubation the tubes were cooled to room temperature and the absorbance at 562nm of each solution determined relative to a water blank. A standard curve was generated from the 562nm absorbance of the known BSA concentration samples and the protein content of each of the cell or tissue lysates determined by comparison to the standard curve.

#### **2.4.3 Protein electrophoresis**

0.3 volumes of 4x SDS-Sample Loading buffer (Novex, see appendix 1) were added to the quantified protein samples. The samples were heated to 70°C for 15 minutes and then vortexed briefly. SDS-polyacrylamide gel electrophoresis was accomplished on precast 10% Tris-Glycine gels in a XCell II™ tank (Novex). A volume of sample in loading buffer containing 20µg of total protein (5-25µl) was added to each of the wells and electrophoresed at 150V in parallel with a molecular weight marker (Kaleidoscope Prestained Standards, BioRad) in MOPS SDS running buffer (Novex, see appendix 1) at room temperature for 1 hour.

#### **2.4.4 Protein Transfer**

Protein was transferred onto Hybond<sup>TM</sup>ECL<sup>TM</sup> nitrocellulose membranes (Amersham) via electrophoresis at a constant current of 250mA per gel for 90 minutes in transfer buffer (see appendix 1). Transfer was carried out at 4°C in a BioRad wet blotting apparatus. Following transfer gels were stained in GelCode (Pierce) to determine transfer efficiency and to check loading levels in the individual wells. Gels were placed in 15 ml of stain solution in a Coplin jar and gentle rocked for 1-2hours then destained in dH<sub>2</sub>O for 2-6 hours. Destained gels were place on 3MM Whatman filter paper and dried in a vacuum apparatus.

#### **2.4.5 Antibody incubation**

Protein blots were transferred to 50ml centrifuge tubes with 10ml of blocking solution (10% non-fat dried milk powder (w/v) in TBST, see appendix 1) and placed on a roller mixer. Blots were blocked in this solution for 1 hour at room temperature. Following blocking, 5 $\mu$ l of primary antibody solution (anti- $\beta$ -catenin, Upstate Biotechnologies or Transduction Laboratories) was added to the blocking solution and the blots incubated in the same manner for 1 hour at room temperature or overnight at 4°C. The primary antibody solution was then poured off and the blot washed 4 times in 10ml TBST at room temperature. The blot was then incubated with 5 $\mu$ l of HRP-conjugate secondary antibody solution (anti-light chain IgG, name of company) in 10ml of 10% Marvel (Nestlé, UK) in TBST (see appendix 1) at room temperature for 1 hour. The blot was washed as above prior to signal detection.

#### **2.4.6 Signal detection**

Signals were detected using the ECL<sup>TM</sup> Chemiluminescent detection system (Amersham). For each blot 1ml of each of the two reagents solutions were mixed just prior to application. The blot was placed on an acetate transparency, protein upwards, and the mixed reagent solution pipetted onto its surface. The reagents were left in contact with the blot for 1 minute then poured off. A second acetate transparency was laid on top of the blot and this sandwich then placed in a film cassette. The blot was then overlaid with ECL<sup>TM</sup> film (Amersham) for periods of 30 seconds to 10 minutes, depending on signal intensity. Films were developed (Hyperprocessor, Amersham), digitally scanned (DeskScan II, Hewlett Packard) and bands quantified with densitometry software (Aida, Microsoft).

## **2.7 Introduction of DNA into cultured mammalian cells**

### **2.7.1 Microinjection**

#### **2.7.1.1 DNA preparations**

DNA for microinjection was diluted to 125ng/ $\mu$ l in 1x injection buffer (see appendix 1) to a final volume of 10 $\mu$ l. The DNA dilutions were spun at 10,000g for 3 minutes to pellet any debris which may clog the injection needle. 2 $\mu$ l of the DNA/buffer solution was placed into the tip of the microinjection needles (Original Femtotips, Eppendorf) using drawn loading pipette tips (Eppendorf).

#### **2.7.1.2 Injection conditions**

Microinjections were carried out on a Zeiss Axiovert S100 Microscope using an Injectman control console and a Transjector Injection system (both from Eppendorf). Injection conditions for the cell lines use were as follows:

**Table 2.4: Microinjection conditions**

<b>Cell line</b>	<b>Injection pressure</b>	<b>Injection time</b>
HCT 116	130 hPa	0.3 seconds
Primary murine colonocytes	100 hPa	0.1 seconds

### **2.7.2 Transfection**

#### **2.7.2.1 Calcium phosphate transfection**

Transfection was carried out using a method described in Maniatis (16.33). Plasmids to be transfected into mammalian cells were linearised with an appropriate restriction enzyme digest followed by phenol: chloroform extraction. The precipitated plasmid DNA was resuspended in the appropriate amount of sterile 0.1x TE (pH 8.0) to give a concentration of 40 $\mu$ g/ml.

Cell lines were plated in 60mm petri dishes 12 to 24 hours prior to transfection. Plasmid DNA was prepared to give a solution volume of 220 $\mu$ l at 40 $\mu$ g/ml and 250 $\mu$ l of 2 x HBS were added. 31  $\mu$ l of 2M CaCl<sub>2</sub> was slowly added to this



solution during continuous but gentle vortexing. The resulting solution was incubated at room temperature for 25 minutes or until a faint precipitate could be seen. 0.5ml of this DNA solution was added drop-wise to each cell culture dish containing 1ml of medium with continuous gentle agitation. 4ml of prewarmed media was then added to each dish and the cells allowed to incubate for 24 hours. After this period the medium was removed, the cells washed in PBS and re-fed with 5ml of media. For selection of G418 resistant clones the cells were trypsinised and replated at low densities and media containing the appropriate concentration of G418 (see table 2.5). The cells were kept under G418 selection for up to 14 days.

**Table 2.5: G418 selection concentrations**

Cell Line	Predetermined lethal G418 Concentration ( $\mu\text{g/ml}$ )
SW480	400
HT29	800
DLD-1	400
HCT116	800

### 2.7.2.2 Liposome mediated transfection

Optimal DNA ( $\mu\text{g}$ ) to FuGene ( $\mu\text{l}$ ) transfection reagent (Roche) ratios for the individual cell lines used in transfection experiments were determined by transfection with the positive control pCMV-GFP reporter plasmid. DNA:reagent ratios of 1:6, 1:3 and 2:3 (recommended by the manufacturer) were used initially with 0.3 $\mu\text{g}$  of DNA being applied to a 75% confluent well of a 24-well plate (Nunc or Grenier). Transfection solutions were prepared by placing the appropriate amount (3 or 6 $\mu\text{l}$ ) of reagent into room temperature media to a volume of 100 $\mu\text{l}$ . 1 or 2 $\mu\text{g}$  of plasmid DNA was added to the FuGene/medium solution and gently mixed followed by incubation at room temperature for 30 minutes. Following the incubation 30 $\mu\text{l}$  (for 1:6 or 1:3 ratios) or 15 $\mu\text{l}$  (for 2:3 ratio) of the solution was added with shaking to each well (containing 1ml of medium). Further experiments varied the amount of transfection solution, and therefore amount of plasmid DNA, which was added to each well. Optimal

transfections were obtained using a 1:6 ratio and 1µg of DNA for cell lines, 2µg of DNA for primary colonocytes and 1µg for murine ES cells.

## **2.8 Animal work**

### **2.8.1 Tissue processing**

Animals showing signs of illness were sacrificed by cervical dislocation and immediately dissected. Organs collected (kidney, liver, spleen, mammary gland and pancreas) were placed immediately in Formalin (see appendix 1), wax embedded, sectioned to 3µm and haematoxylin and eosin stained. Intestines were dissected out of the animals intact and flushed with water to remove waste material. The large and small intestines were then separated, placed on Whatman 3MM paper and cut open along their length. The opened intestines, still on 3MM filter paper, were placed in a tray containing Methocarn (see appendix 1) and left to fix for 3-12 hours. The intestine samples were then “swiss-rolled” by winding onto the end of watchmaker forceps and skewered by a syringe needle. These gut rolls were submerged in Formalin (see appendix 1), wax mounted, sectioned to 2µm and stained with haematoxylin and eosin. Tissue sections were visually examined for the presence of neoplastic lesions and/or other histological abnormalities.

### **2.8.2 Timed matings**

For the harvest of embryos for prenatal phenotype examination, virgin females were caged with males and were examined daily to check for the presence of a post-coital (pc) plug. On the appropriate day (8-10 days pc) the pregnant females were sacrificed by cervical dislocation and the uterus immediately dissected out and placed in ice-cold PBS. Dissection of embryos from the maternal tissue and placenta was carried out in a petri dish containing ice-cold PBS. Yolk sac samples were retained for DNA extraction and genotyping of individual embryos where appropriate. Embryos were examined under a binocular dissection microscope (Heerbrug) for any gross anatomical abnormalities and to more accurately establish developmental stage. Embryos were then fixed in formalin (see appendix 1), wax embedded and sectioned.

# **Chapter 3**

## **Time and tissue-specific gene deletion**

## Chapter 3 - Time and tissue-specific gene deletion

### 3.1 Background

While the *Min* mouse model and other murine APC mutants, such as the APC 1638N line, have provided considerable detail regarding the genetic progression from normal tissues to neoplastic colorectal tumours these models have certain limitations (Smits *et al.*, 1998). As discussed in chapter 1, the heterozygous loss of APC function in all tissue types during the entire lifetime of a mouse does generate an animal model of tumour formation similar to the rare human condition FAP, but does not adequately model the more common condition of sporadic colorectal carcinoma. Tumour formation in the mouse is predominantly in the small intestine, while in the human small intestinal tumours are very rare compared to tumours of the large intestine. In order to better understand the formation of sporadic colorectal lesions it would be useful to possess an animal model that more closely resembles the human phenotype. One way to achieve this would be to be able to control more precisely where and when the murine APC gene is lost. The aim of this project was to generate an intestinal stem cell specific inducible Cre mouse strain, with the intention of crossing it to a transgenic mouse strain containing one or more copies of a floxed wildtype *Apc* gene compensating for homozygous *Min* mutations. On induction of Cre recombinase activity, the floxed *Apc* genes would be excised only in the stem cells and proliferative compartment, hopefully resulting in a phenotype resembling that observed in the human sporadic colorectal cancer patient. A floxed *Apc* / *Apc*<sup>Min/Min</sup> mouse strain has been developed and it has been shown that introduction of Cre coding sequence into the intestinal epithelial cells of these mice using engineered adenovirus results in rapid formation of polyps and eventually adenomas (see section 1.5.3.6.2) (Shibata *et al.*, 1997). However, as the Cre-adenovirus does not only infect stem cells, some of the polyps or adenomas may arise from older cells which have accumulated mutations capable of influencing tumour formation and progression. Additionally, the number of tumours formed following adenovirus Cre delivery was very small when compared with the number of cells infected using a related adenovirus encoding the LacZ gene. The difference in the frequency of polyp formation compared with the frequency of adenovirus infection in the intestinal epithelium suggests

that either recombination is not occurring in an efficient manner or that loss of *Apc* function alone is insufficient to cause polyp formation. While it is known that most intestinal tumours contain alterations to both copies of the *Apc* gene and that early adenomas may be wildtype for a number of other tumour suppressor genes, the use of adenovirus Cre delivery does not answer the question of whether these *Apc* mutations alone are sufficient to initiate and maintain tumour formation (Lamlum *et al.*, 2000).

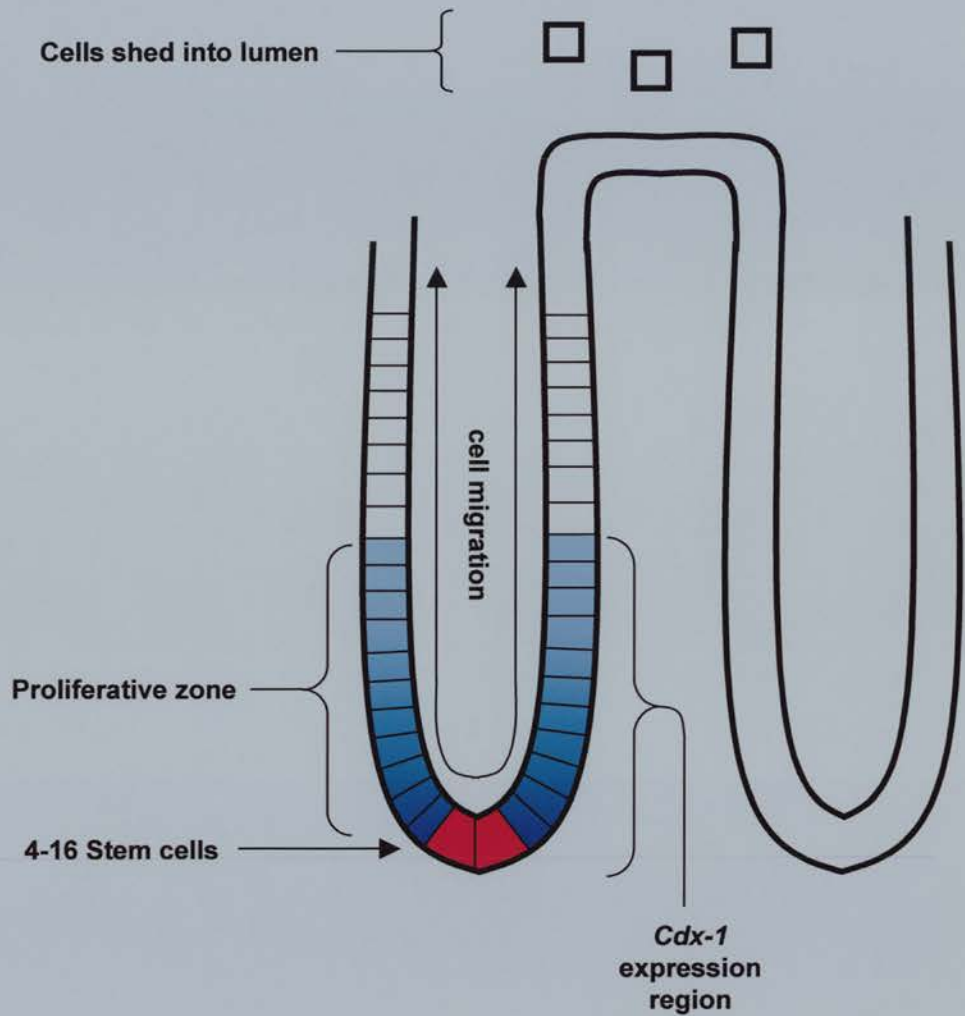
Mouse strains have previously been developed which express the Cre recombinase in the epithelium of the intestine in both inducible and constitutive manners (Saam and Gordon, 1999). In the constitutive expression mouse line the control of Cre transcription is placed under the control of the Fatty Acid Binding Protein (*Fabp*); expression of Cre from this promoter appears to be limited to the epithelial cells of most of the length of the mouse small intestine. In the inducible recombination mouse the reverse tetracycline (tet) transactivator sequence is under the control of the *Fabp* promoter while Cre expression occurs from a minimal human cytomegalovirus (CMV) promoter, itself under the control of a tet operator sequence, at another transgenic locus. In this system, expression of the tet transactivator is limited to the small intestinal epithelium; administration of the tetracycline analogue doxycycline activates this transcription factor which then binds to the tet operator and provokes expression of the Cre coding sequence, thereby also limiting Cre expression to the small intestinal epithelium. Analysis of these mouse strains crossed onto floxed reporter strains indicate that when Cre is under the constitutive control of the *Fabp* promoter, Cre recombination can be observed prenatally from day 13.5 post-conception while in the inducible mouse strain recombination occurs with high efficiency after administration of doxycycline. The presence of the recombined allele is detectable up to 60 days after the cessation of doxycycline treatment despite the high cell turn-over rate of the intestinal epithelium, indicating that the *Fabp* promoter is active in cells with proliferative potential (Saam and Gordon, 1999). While this model will prove useful for the investigation of the effects of gene removal in the developing (in the case of the constitutively expressed Cre) and in the adult (in the inducible Cre mouse) intestine, it has not been clearly established whether the *Fabp* promoter initiates expression in the stem cells of

the intestinal crypts. The ability to induce Cre recombination in the crypt stem cells would allow for the expansion of clonal populations derived from the altered progenitor cells.

In order to generate such a model organism, tissue-specific promoter elements are required to target the inducible Cre expression to the desired cells. No intestinal stem cell specific promoters have been identified to date; the closest candidate gene proved to be *Cdx-1*, a murine homeobox gene, which is expressed biphasically in the lifetime of the mouse (Meyer and Gruss, 1993). *Cdx-1* is first expressed in the mouse embryo prior to the gastrulation stage (~day 7.5) in the primitive streak, with expression becoming more global later in embryogenesis. In the adult mouse *Cdx-1* expression is limited to the intestinal epithelium, in a gradient increasing from the proximal small intestine to the distal large intestine and is expressed in detectable levels in the stem cells and proliferative compartments of the crypts (see figure 3.1) (Subramanian *et al.*, 1998). The early embryonic expression of *Cdx-1* would prove to be a problem if used to drive native Cre expression in a floxed APC cross line, as APC null homozygosity is embryonic lethal and premature excision of the APC gene in all tissues as a result of Cre activation would result in embryo death (Moser *et al.*, 1995). To avoid the premature excision of APC in this crossed line, the Cre coding sequence is fused to a mutated human oestrogen-binding (ER<sup>tm</sup>) element coding sequence, the resulting fusion protein termed CreER<sup>tm</sup>. This element is normally bound to a heat shock protein thereby sterically hindering the active site of the Cre recombinase, but will recognise and bind the synthetic hormone tamoxifen (<sup>tm</sup>); binding of tamoxifen displaces the heat shock protein from the oestrogen-binding element, thereby restoring Cre recombinase activity (Feil *et al.*, 1996). The combination of the tissue-specific *Cdx-1* promoter elements with the tamoxifen inducible Cre would allow control of APC excision in a time and tissue dependant manner, allowing for better modelling of the formation of human sporadic colorectal adenomas and the investigation of whether APC inactivation alone is sufficient to cause tumour formation.



**Figure 3.1: Expression pattern of *Cdx-1* in the murine colonic crypt**



**Figure 3.1:** *Cdx-1* expression is limited, in the adult mouse large intestine, to the proliferative compartment and stem cells of the colonic crypts. *Cdx-1* expression is also detected at lower levels in the small intestine, again in the proliferative compartment.

## 3.2 Isolation of *Cdx-1* tissue-specific promoter sequence

### 3.2.1 BAC screening

The initial step in generating an intestinal specific inducible Cre mouse line was the identification and cloning of the *Cdx-1* promoter elements needed to confer tissue-specific promoter activity. At the time of the commencement of this project approximately 1.1kb of sequence upstream of the transcription start site of *Cdx-1* had been published and was available from GenBank (Hu *et al.*, 1993). This short stretch of sequence had been engineered into a construct to drive expression of a Chloramphenicol Acetyltransferase (CAT) reporter gene when transiently transfected into a number of colorectal adenoma derived cell lines (Hu *et al.*, 1993). Levels of transcription generated using this short promoter sequence were extremely low and only produced measurable levels of reporter protein when a strong viral enhancer (SV40 enhancer) was present in the reporter construct (Hu *et al.*, 1993). This 1.1kbp of published *Cdx-1* promoter sequence did not appear able to drive significant levels of reporter transcription in the absence of an exogenous enhancer and did not appear to contain an endogenous enhancer, although it was found to contain an orientation independent silencer element and a silencer blocking element. (Hu *et al.*, 1993). While the published sequence did contain a number of transcriptional control elements, it was still possible that the published *Cdx-1* promoter did not contain all the elements necessary to confer the precise tissue-specificity required; transcription factor binding sites might exist upstream of this initial 1.1 kb of promoter sequence which are vital to producing not only reasonable levels of transcription but the observed expression pattern.

The first step in generating constructs driven by the *Cdx-1* promoter was therefore the isolation of a significant amount of sequence upstream of the *Cdx-1* transcription start site. A number of methods for isolating sequence immediately upstream of the published *Cdx-1* promoter sequence were investigated, including genome walking and inverse PCR. While some sequence was successfully isolated and sequenced using an inverse PCR method (see figure 3.2) the lack of suitable restriction sites near the 5' limit of the known *Cdx-1* sequence limited the viability of finding large amounts of new sequence. Screening available

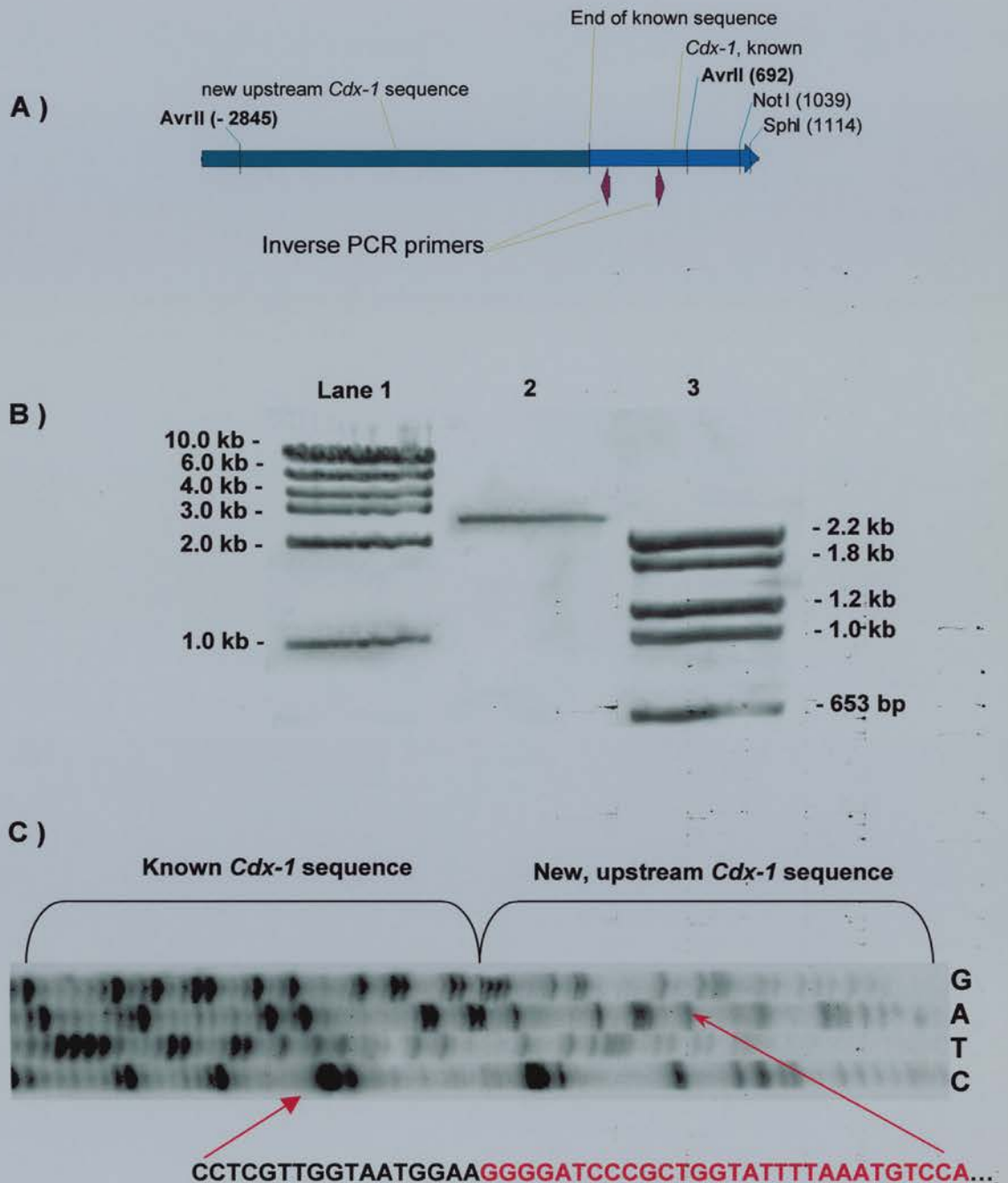
genome libraries using probes based on the published *Cdx-1* sequence was determined to be the most reliable way to isolate the desired elements. Ultimately, it was decided to screen a Bacterial Artificial Chromosome (BAC) library to isolate clones containing murine *Cdx-1* sequence (Shizuya *et al.*, 1992). The very large insert sizes in these pBeloBAC11 constructs (>100kb) would maximise the chances of the known *Cdx-1* sequence being contained in the same clone as the unknown upstream sequence of the *Cdx-1* promoter region (Monaco and Larin, 1994). The initial stages of BAC screening involve PCR reactions; BAC DNA was obtained from Research Genetics and screened by PCR for the presence of exon 1 of *Cdx-1* using primers and conditions optimised by Dr. Jane Armstrong (figure 3.3, see section 2.2.2 for PCR conditions). The genomic mouse DNA inserts in these BAC constructs were derived from the CJ73 cell line which was derived from the 129SV mouse strain. DNA samples were organised in superpools of DNA derived from a large number of clones and pools containing a much smaller number of BAC DNA samples. Once the correct pool of BAC DNA had been identified a membrane corresponding to this DNA sample was obtained for the final round of BAC library screening. This membrane contained DNA dot blots from 384 BAC host clones and was hybridised to a PCR amplified *Cdx-1* exon 1 DNA probe by Dr. Sula Corbet. Two BAC clones were identified as positive for the presence of *Cdx-1* exon 1, BACs 219 and 538 (see figure 3.4).

### **3.3 Construction of BAC based targeting and transgenic vectors**

In order to maximize the amount of upstream *Cdx-1* sequence included in the transgene construct it was decided to employ a relatively recently developed technology to build a *Cdx-1* driven CreER<sup>tm</sup> transgene in parallel with the more traditional subcloning methods. A method for the rapid modification of BAC sequence based on the viral Red $\alpha$ - $\beta$ - $\gamma$  recombination-inducing proteins has been used to introduce changes into specific locations in BACs without the need to purify large quantities of BAC DNA or indeed remove the BAC from its host bacteria (Muyrers *et al.*, 1999). The method employs an inducible Red $\alpha$ - $\beta$ - $\gamma$  expression plasmid, which is first introduced into the BAC host organism by

electroporation (see section 2.1.7). The plasmid is induced by the addition of arabinose to the culture media which results in the high level expression of Red $\alpha$  to complement the constitutive expression of Red $\beta$ /Red $\gamma$  and these “activated” recombination competent cells are harvested and made electrocompetent (see figure 3.5, section 2.1.7). The recombination competent BAC host bacteria are then electroporated with a linearized vector containing sequence to be inserted into the BAC along with a selection cassette to screen for bacteria containing recombinant BACs. This sequence to be introduced is flanked with short stretches of DNA (40-150bp) homologous to the region of the BAC to be targeted for insertion (see figure 3.6 for overview). The recombination-induction of the Red $\alpha$ - $\beta$ - $\gamma$  proteins and the ability to select for recombinants using an antibiotic system should make this process very efficient.

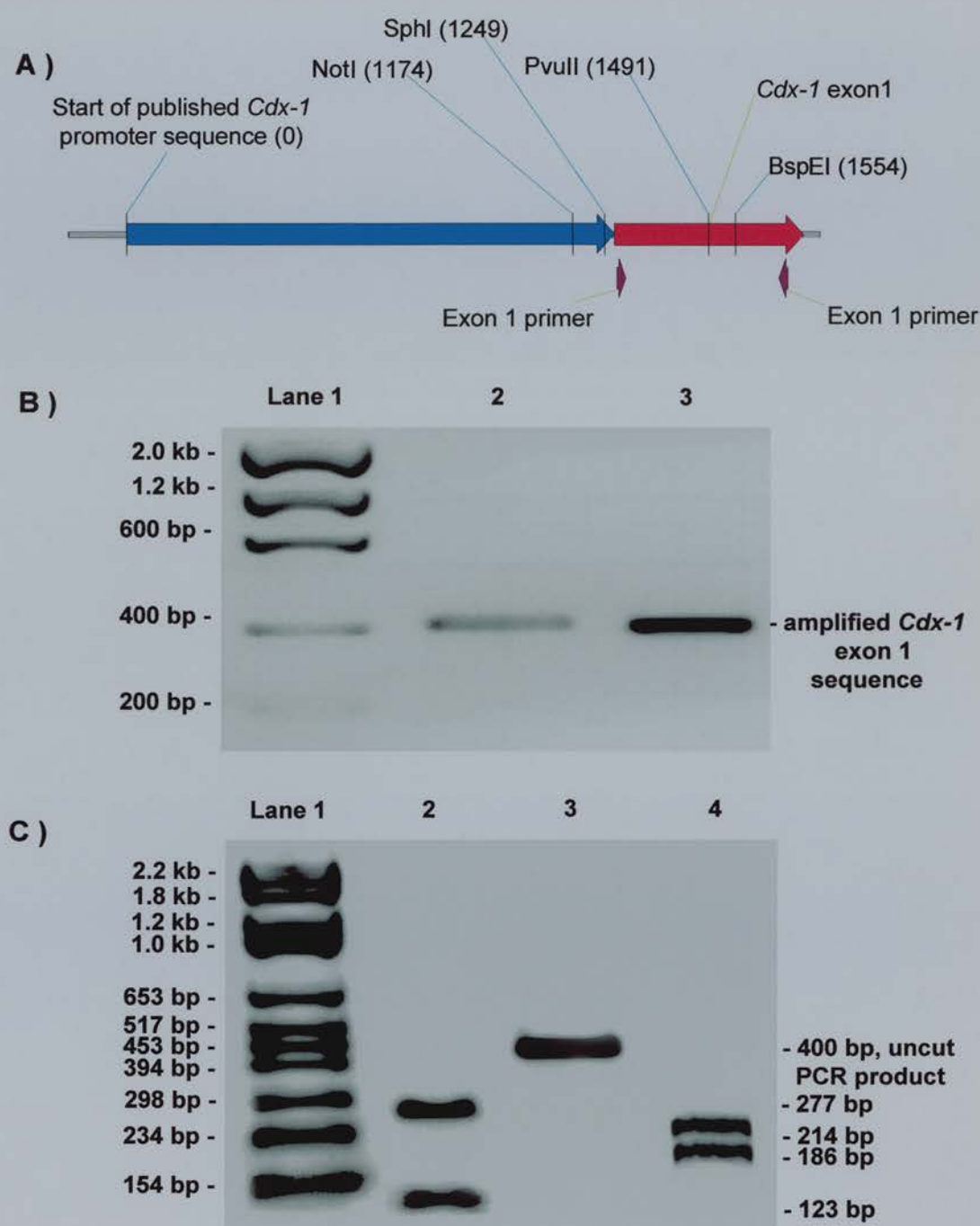
**Figure 3.2: Inverse PCR for the isolation of upstream *Cdx-1* promoter sequence**



**Figure 3.2:** **A)** Schematic of the region amplified in the inverse PCR of circularised *AvrII* digested mouse genomic DNA showing the locations of the PCR primers as well as the subsequently identified location of the upstream *AvrII* site. **B)** UV gel image of a representative PCR product generated from the inverse template. Lane 1: High Mass Marker (Gibco). Lane 2: Inverse PCR product of 2.8 kb generated from *AvrII* digested/circularised mouse genomic DNA. Lane 3: Marker VI (Roche). **C)** Sequence derived from the pCR 2.1 cloned inverse PCR product.

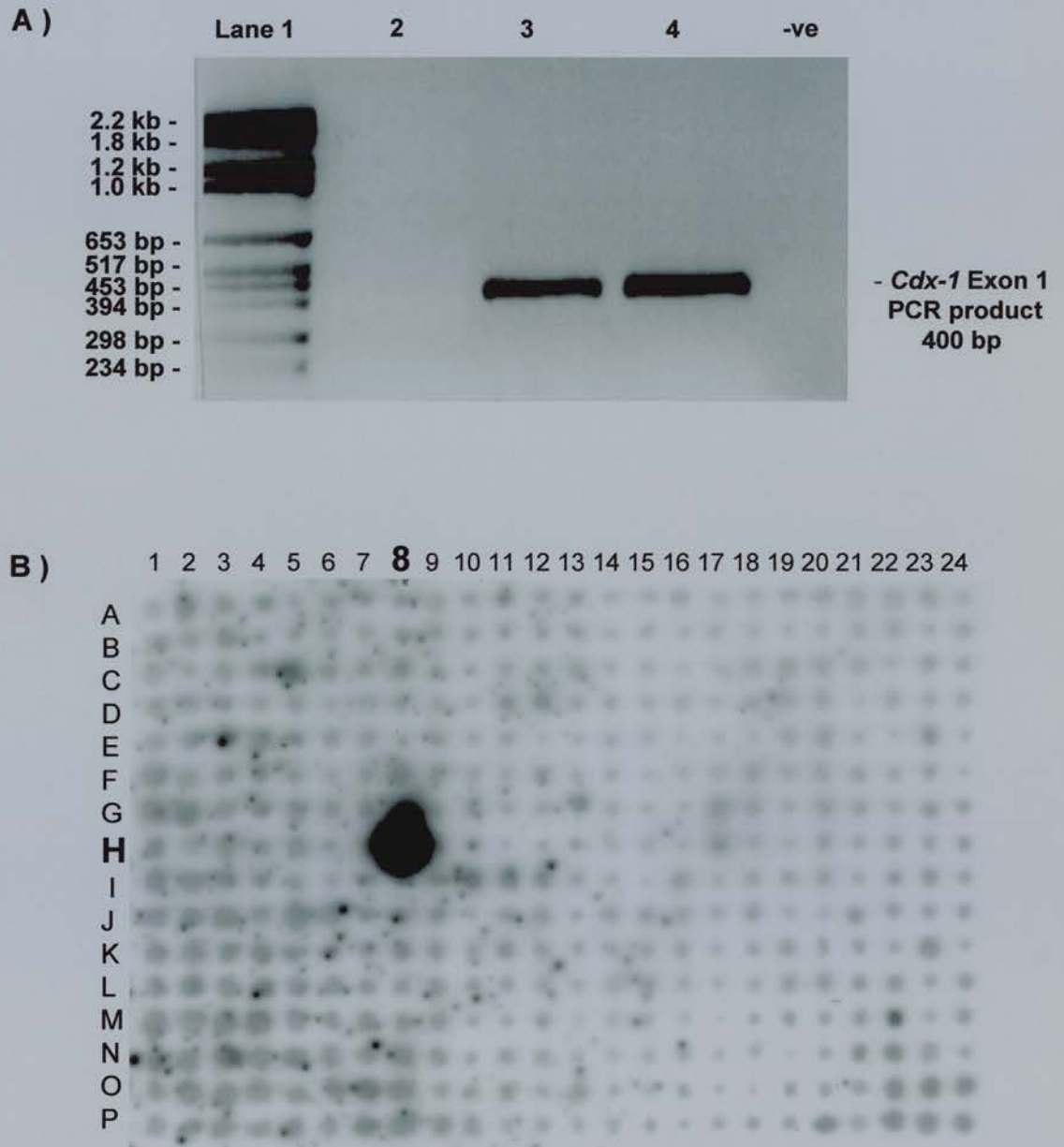


**Figure 3.3: *Cdx-1* exon1 PCR for screening of BAC library**



**Figure 3.3:** **A)** Schematic of the published *Cdx-1* promoter region and exon1, showing some relevant restriction sites and the location of primers used to amplify the exonic sequence. **B)** Gel electrophoresis of representative PCR reaction of the first exon of *Cdx-1*. Lane 1: Low Mass Ladder (Gibco). Lane 2: PCR product produced in a 25µl reaction. Lane 3: PCR product produced in a 50µl reaction. **C)** Analytic digests of the PCR product to confirm the correct amplification of exon 1 of *Cdx-1*. Lane 1: Marker VI (Roche). Lane 2: BspEI digestion of the amplified 400 bp exon 1 sequence showing the expected bands at 123 and 277 bp. Lane 3: Undigested exon 1 PCR product. Lane 4: PvuII digestion of the amplified 400 bp exon 1 sequence showing the expected bands at 186 and 214 bp.

**Figure 3.4: Isolation of *Cdx-1* positive BAC clone**



**Figure 3.4:** **A)** A representative 400 bp PCR reaction screening for the presence of the first exon (exon 1) of the murine *Cdx-1* gene within the BAC DNA pools provided. Lane 1: Marker VI (Roche). Lane 2: *Cdx-1* exon 1 PCR of negative BAC DNA pool 2F. Lanes 3-4: Positive DNA pools 8B and 4C respectively. Lane 5: A negative control reaction using dH<sub>2</sub>O in the place of template DNA. **B)** An autoradiograph of a membrane dot blotted with the DNA from BAC clones making up the positive DNA pool. The membrane was hybridized with a PCR amplified 400 bp stretch of exon 1 of murine *Cdx-1*. The positive clone gives a strong signal and is found at position H8, corresponding, in this case, to BAC clone 219.

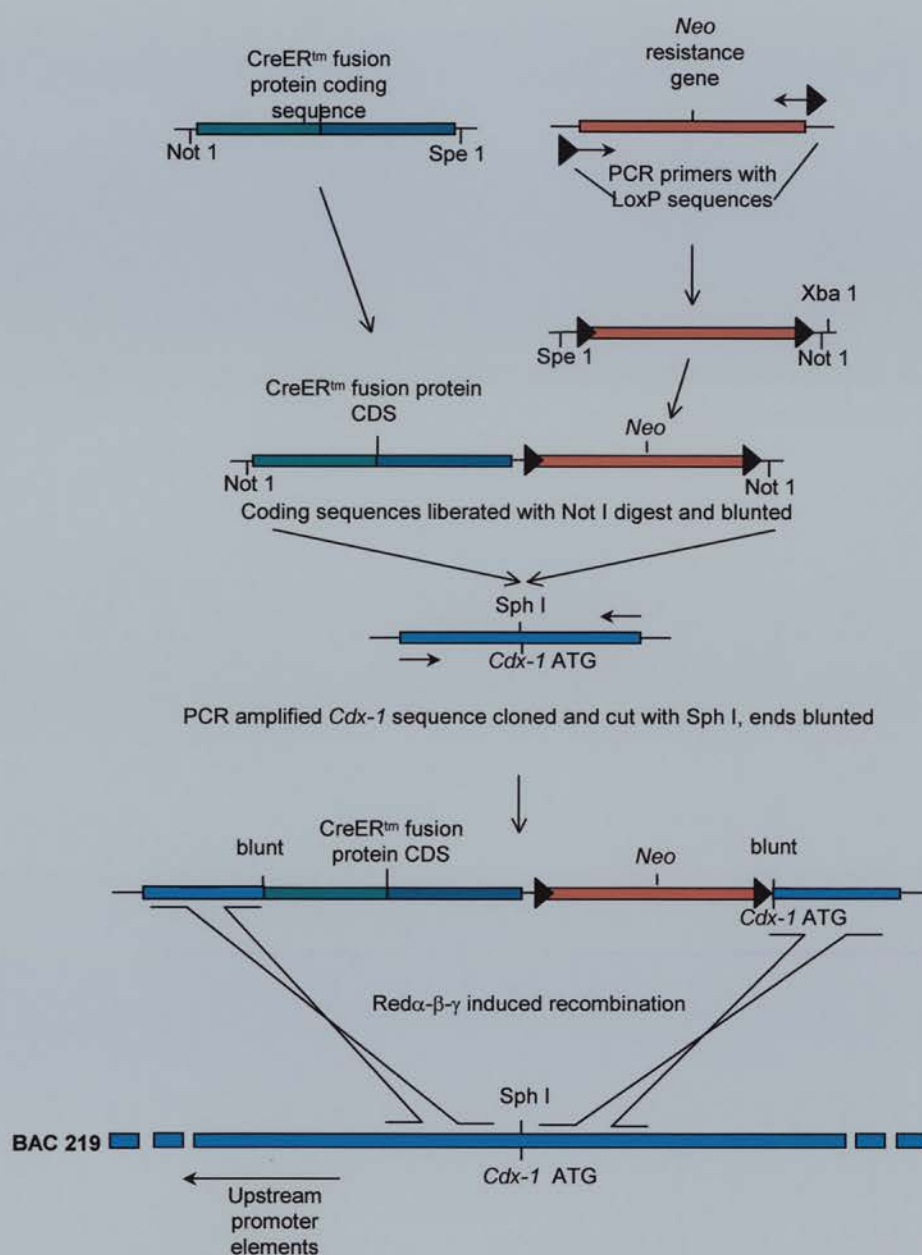


**Figure 3.5: Induction of the Red $\alpha$  protein in bacteria by exposure to arabinose**



**Figure 3.5:** A GelCode (Pierce) stained SDS-PAGE electrophoresis gel of total proteins isolated from bacterial host cells containing the pBAD-Red $\alpha$ / $\beta$ / $\gamma$  recombination inducing plasmid. The expression of the Red $\alpha$  protein is under the control of an arabinose inducible promoter (Clontech). Lane 1: Broad Range Pre-Stained Protein Standards (BioRad). Lane 2: Total proteins isolated from BAC 219 host *E. coli* cells cultured in the absence of arabinose. Lane 3: Total proteins isolated from BAC 219 host *E. coli* cells cultured in LB media supplemented with 0.2% w/v arabinose showing the induction of the 25 kDa Red $\alpha$  protein.

**Figure 3.6: Strategy for the creation of a BAC based targeting construct**

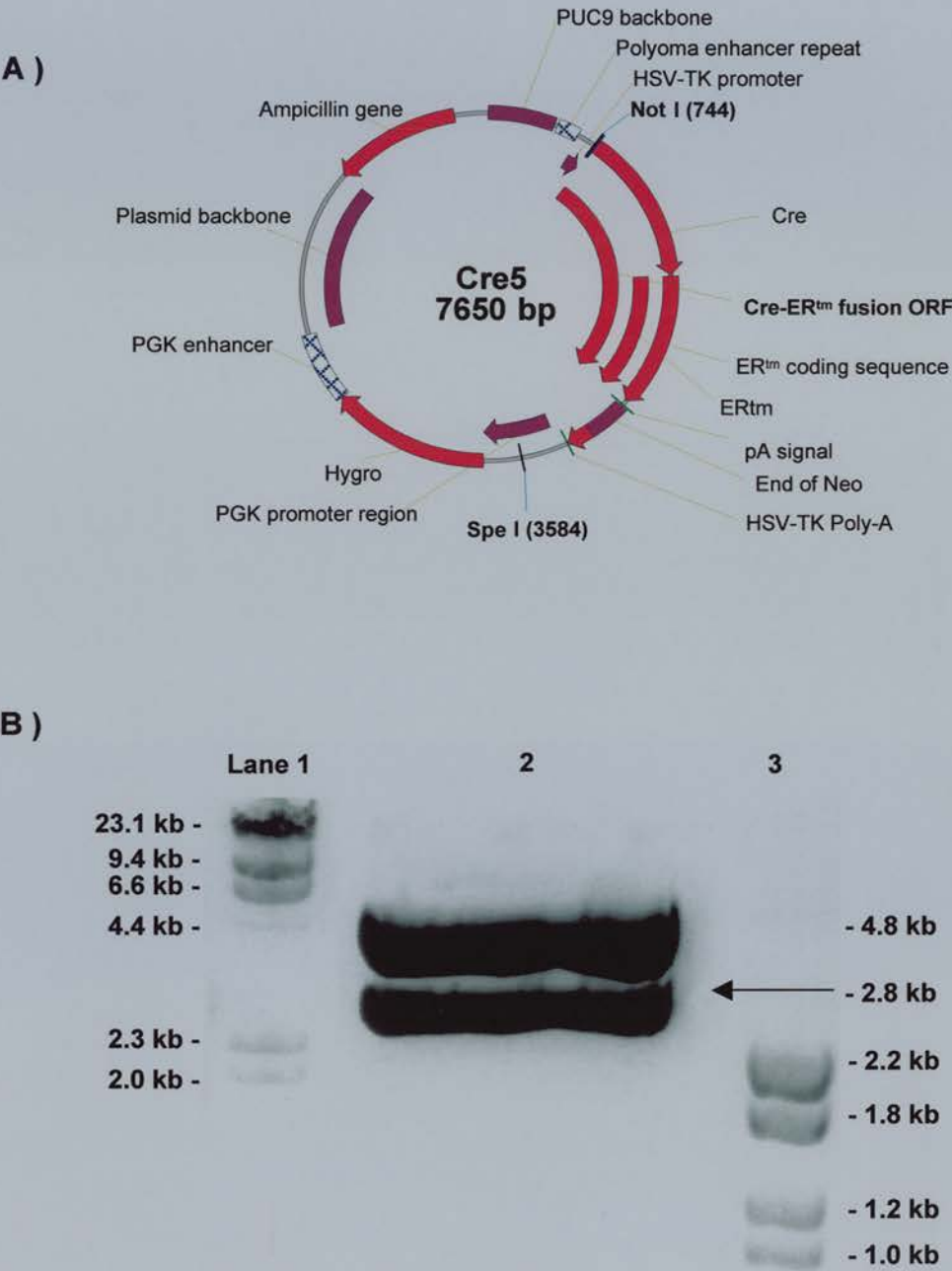


**Figure 3.6:** An overview of the strategy for the creation of a BAC based gene targeting construct for the expression of inducible Cre recombinase under *Cdx-1* promoter control using the Red  $\alpha/\beta$  recombination technique. Floxed *Neo* resistance cassettes were cloned into a vector containing the promoterless CreER<sup>tm</sup> fusion protein coding sequence. The CreER<sup>tm</sup> and resistance cassette were excised from the plasmid backbone and inserted into the sequence of a short stretch of cloned *Cdx-1* sequence containing the transcription start site. The resulting plasmid was linearised and transfected into bacteria containing a Bacterial Artificial Chromosome (BAC) encoding the entire murine *Cdx-1* promoter sequence and an activated Red  $\alpha/\beta$  plasmid. The CreER<sup>tm</sup> Neo plasmid underwent recombination with the wild type sequence of the BAC producing a Cre expression construct under the control of the entire *Cdx-1* promoter. To create an analogous transgene construct the *Neo* gene was not flanked by loxP sites but the strategy remained the same.

### 3.3.1 Subcloning of CreER<sup>tm</sup> coding sequence

The coding sequence of the Cre recombinase fused to that of a human oestrogen receptor altered to preferentially bind the synthetic hormone tamoxifen, CreER<sup>tm</sup>, was obtained from Dr. Tom Gardner. This plasmid, termed Cre5 (see figure 3.7a), contained a number of stretches of sequence which would have interfered with the construction or function of the final BAC targeting vector, making it necessary to subclone the coding sequence into another vector. The CreER<sup>tm</sup> coding sequence was excised from the Cre5 plasmid by a double digestion with SpeI and NotI, followed by gel extraction of the 2840 bp fragment containing the CreER<sup>tm</sup> sequence (figure 3.7b). This fragment was cloned into the pBlueScript (pBS) (Stratagene) vector, which had been digested with the same enzymes and treated with SAP to prevent recircularisation. The result was a “clean” CreER<sup>tm</sup> coding sequence vector, pBSCre without the viral promoters and stretches of unknown sequence found in the original Cre5 plasmid (figure 3.8).

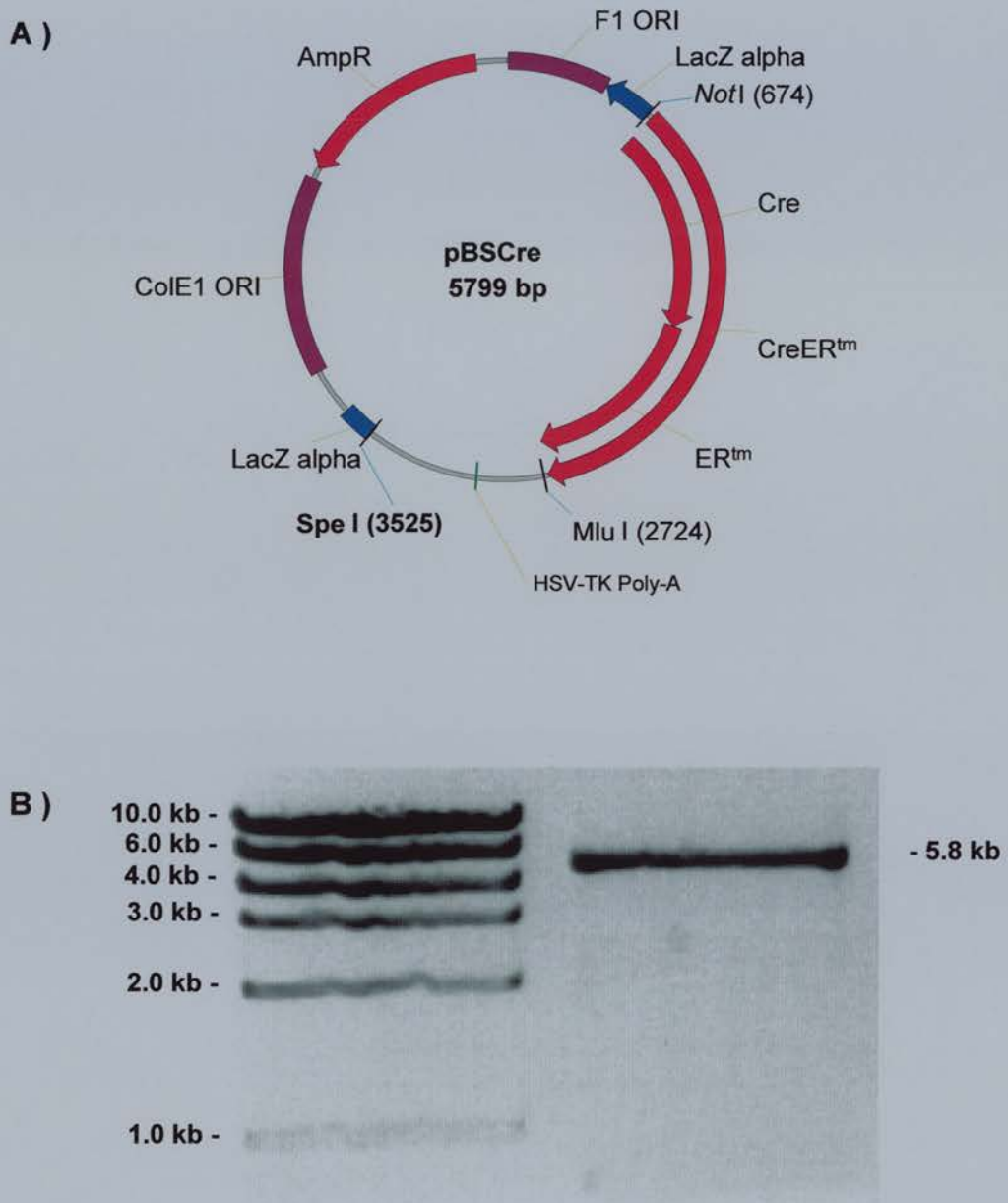
**Figure 3.7: Subcloning of the CreER<sup>tm</sup> coding sequence**



**Figure 3.7: A)** A vector schematic of the Cre5 vector, constructed by Dr. Tom Gardner, coding the CreER<sup>tm</sup> fusion protein. **B)** Electrophoresis gel showing the NotI/ SpeI digest of the Cre5 vector for gel extraction and subcloning of the Cre-ER<sup>tm</sup> fusion ORF into the pBlueScript backbone. Lane 1: Marker 2 (Roche). Lane 2: Cre5 digested with NotI/SpeI. The DNA fragment encoding the CreER<sup>tm</sup> protein is present in the 2.8 kb band indicated.



**Figure 3.8: The pBSCre plasmid**



**Figure 3.8:** **A)** A vector schematic showing the CreER<sup>tm</sup> coding sequence cloned into the pBlueScript (pBS) plasmid backbone. **B)** Gel image of the pBSCre plasmid prepared for the next cloning step. Lane 1: High Mass Marker (gIBCO). Lane 2: pBSCre digested with SpeI giving a single band at 5.8 kb representing the linearised plasmid. The digestion was supplemented with SAP to prevent recircularisation in subsequent cloning stages.

### 3.3.2 PCR generation of mutant floxed and native kanamycin/neomycin resistance cassettes

In addition to the CreER<sup>tm</sup> coding sequence it was necessary to generate an appropriate antibiotic selection cassette. Two uses for the recombination altered BAC were envisioned at the inception of this project; the use of a full length linearised BAC construct as a random integration transgene and, should appropriate “trimming” digests prove possible, as a targeting vector for the replacement of one native copy of *Cdx-1* in murine embryonic stem (ES) cells with the *Cdx-1* promoter driven CreER<sup>tm</sup> recombinase. Because of the need to select for recombinant BACs in prokaryotic hosts as well as possibly needing to select for ES cells which had undergone recombination, the selection cassette chosen for use in the BAC based construct was a *Neo* gene driven by both prokaryotic and eukaryotic (Herpes Simplex Virus-Thymidine Kinase or HSV-TK) promoter sequences. The resulting cassette gives resistance to the antibiotic Kanamycin in the *E. coli* K12 derivative cells used during the cloning stages of construct building and would impart resistance to G418 in ES cells which had successfully integrated the altered BAC (Southern and Berg, 1982).

Two cassettes containing the *Neo* resistance gene were generated by PCR from a commercially available vector (pGFP-N1, Clontech). The first cassette contained only the promoter and resistance coding sequences mentioned above while the second cassette generated was flanked by loxP sites containing a single basepair mutation in the spacer region (see section 2.2.2 for primers). This single base change results in loxP sites which will preferentially recombine with each other rather than with native loxP sites (Lee and Saito, 1998). The inclusion of these loxP sites would allow for the removal of the selection cassette from recombination positive “knock-out” ES cells by transient transfection with a Cre coding plasmid driven by a strong promoter (Abuin and Bradley, 1996). Removal of the selection cassette may prevent read through from the selection cassette promoter into sequence downstream of the targeted gene and additionally stops the constitutive transcription and translation of resistance markers in any animal generated using the targeted ES cells; both situations which may result in phenotype unrelated to the desired genomic alterations.

Mutated loxP sites were used to prevent Cre mediated recombination between the single loxP site remaining in the altered *Cdx-1* allele following selection cassette excision and the native loxP sites flanking genes of interest in mice to which the *Cdx-1*-Cre mouse may be crossed, thereby preventing possible large genomic rearrangements. This loxP flanked selection cassette would, however, not be appropriate for use in the random integration transgenic approach as the presence of multiple sites of integration and/or concatamer integration events could result in multiple undesired recombination events upon Cre induction. The *Neo* resistance cassette without loxP sites was therefore generated for use in the transgenic BAC construct while the mutant loxP cassette was generated for use in the targeting BAC construct.

Resistance cassette amplification reactions, both with and without loxP sites, were incubated with Dpn I, a methylation dependant restriction enzyme, following amplification to cleave any template plasmid present in the PCR reaction. The cassettes were then cloned into the pCR 2.1 vector (Invitrogen), resulting in the plasmids pCRNeo and pCRLoxNeo (see figures 3.9 and 3.10). As the pCR 2.1 cloning vector contains a native Kanamycin selection cassette the PCR generated cassettes were cut out with a SpeI / XbaI double digest and cloned into the pBlueScript vector, which contains only the Ampicillin resistance gene, to test for functionality of the PCR generated *Neo* gene, the resulting vectors referred to as pBSNeo and pBSLoxNeo (figures 3.11 and 3.12). Both the floxed and the non-floxed cassettes successfully conferred kanamycin resistance to *E. coli* host strains.

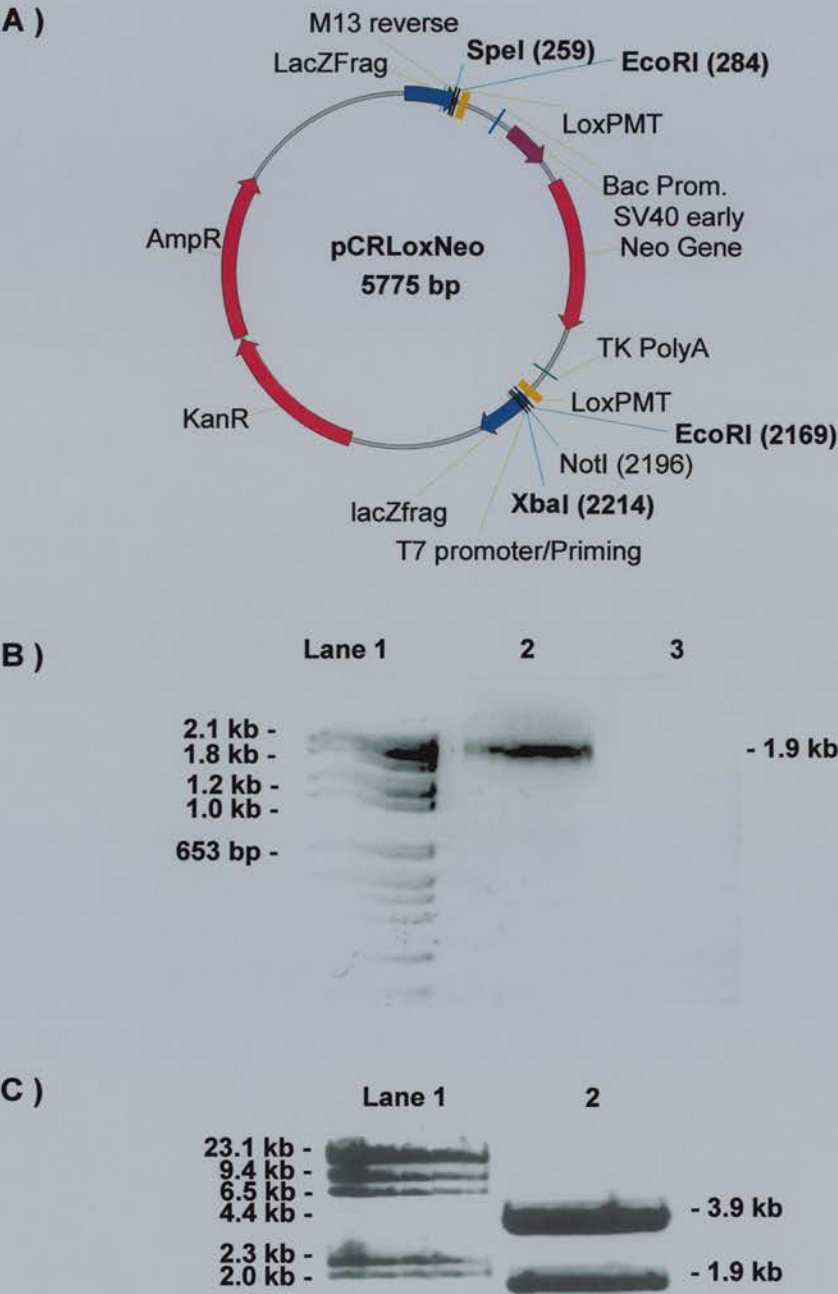


**Figure 3.9: Generation and cloning the Native Neo resistance cassette**



**Figure 3.9:** **A)** A vector schematic showing the non-floxed Neo resistance cassette cloned into the pCR 2.1 plasmid backbone. **B)** A representative PCR product. Lane 1: Kilobase Ladder (New England Biolabs). Lane 2: Negative control using dH<sub>2</sub>O instead of template DNA. Lane 3: 1.9 kb Neo PCR product. **C)** An analytic digest of the cloned Neo PCR product. Lane 1: Marker II (Roche). Lane 2: EcoRI digest of the pCRNeo plasmid. The EcoRI digest liberates the 1.9 kb resistance plasmid from the 3.9 kb pCR backbone. Lane 3: Marker VI (Roche).

**Figure 3.10: Generation and cloning the floxed Neo resistance cassette**



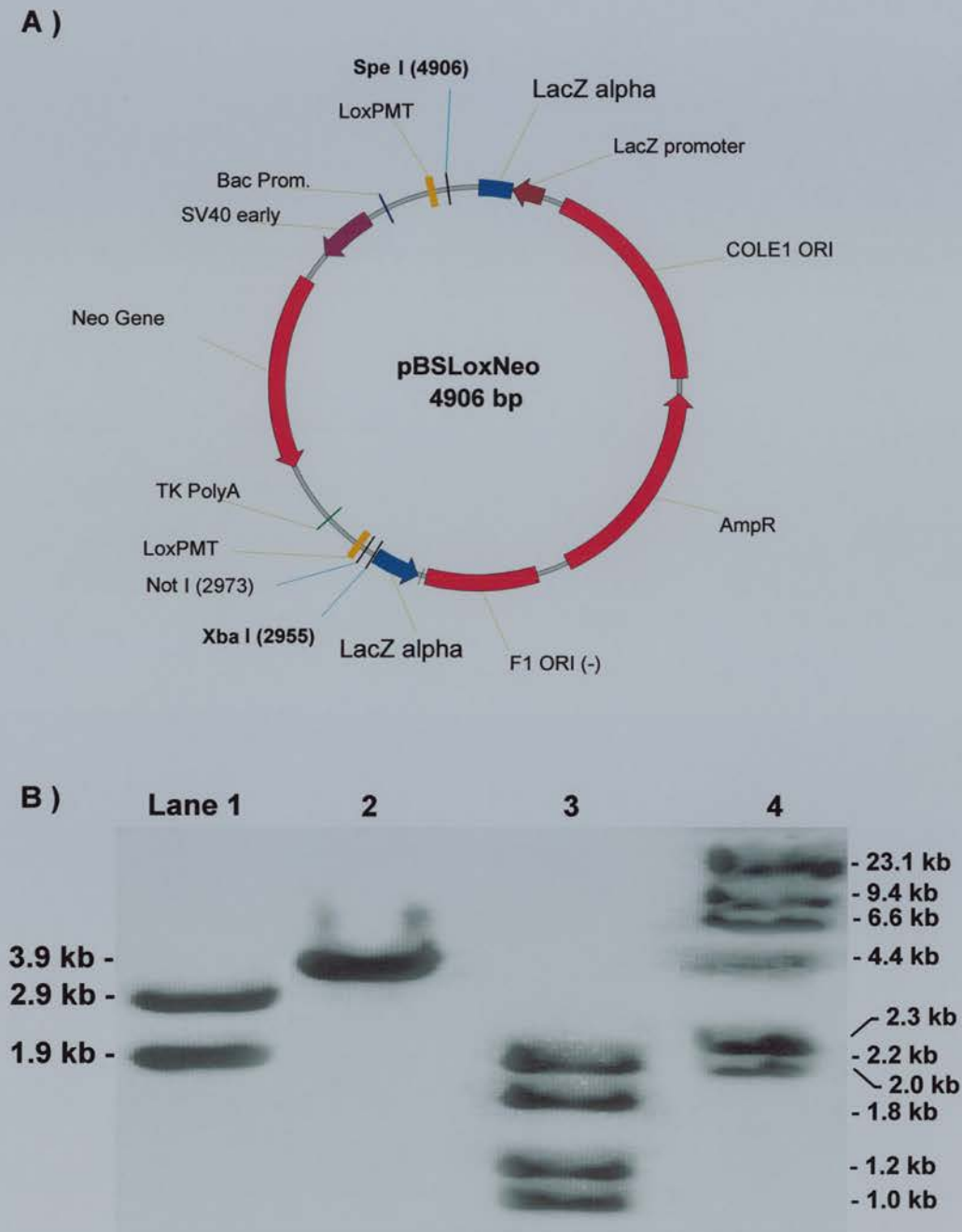
**Figure 3.10:** A) A vector schematic showing the floxed Neo resistance cassette cloned into the pCR 2.1 plasmid backbone. B) A representative PCR product. Lane 1: Marker VI (Roche). Lane 2: 1.9 kb LoxNeo PCR product. Lane 3: Negative control using dH<sub>2</sub>O instead of template DNA. C) An analytic digest of the cloned LoxNeo PCR product. Lane 1: Marker II (Roche). Lane 2: EcoRI digest of the pCRLoxNeo plasmid. The EcoRI digest liberates the 1.9 kb resistance plasmid from the 3.9 kb pCR backbone.

**Figure 3.11: Subcloning of the native Neo cassette**



**Figure 3.11: A)** A vector schematic showing the Neo resistance cassette cloned into the pBlueScript plasmid backbone. **B)** An analytic digest of the correct pBSNeo plasmid. Lane 1: Marker VI (Roche). Lane 2: pCRNeo digested with XbaI/SpeI. This double digestion liberates the Neo cassette from the pCR 2.1 backbone giving bands of 1.9 and 3.9 kb. Lane 3: pBSNeo digested with XbaI/SpeI. This double digestion liberates the Neo cassette from the pBS backbone giving bands of 1.9 and 2.9 kb, showing the successful subcloning of the resistance cassette. Lane 4: Marker II (Roche).

**Figure 3.12: Subcloning of the floxed Neo cassette**



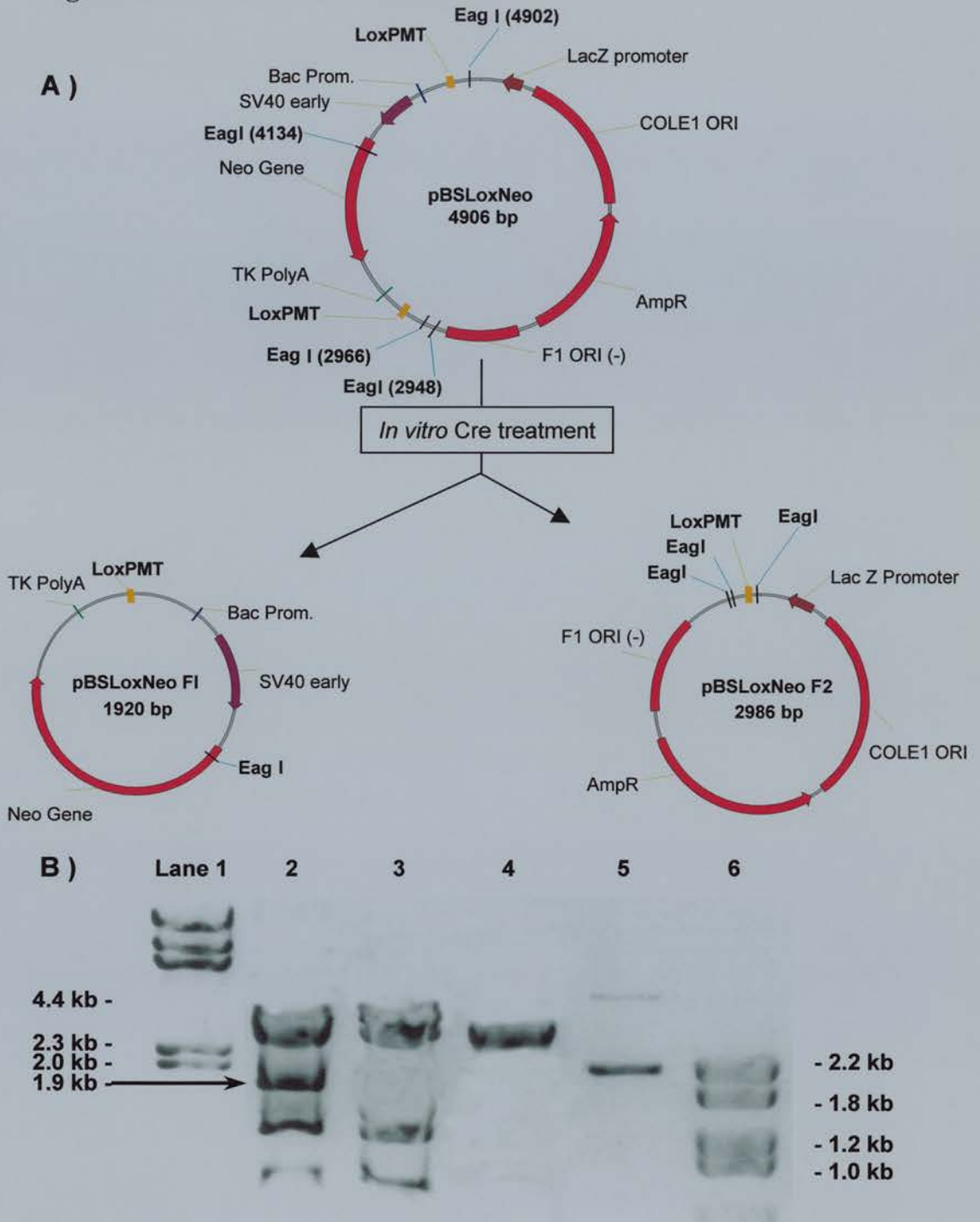
**Figure 3.12: A)** A vector schematic showing the LoxNeo resistance cassette cloned into the pBlueScript plasmid backbone. **B)** A gel image of an analytic digest of the correct pBSLoxNeo plasmid. Lane 1: pBSLoxNeo cut with XbaI/SpeI. This double digestion liberates the 1.9 kb LoxNeo resistance cassette from the 2.9 kb pBS backbone. Lane 2: The 3.9 kb pCR 2.1 vector, from which the LoxNeo fragment had been subcloned, cut with XbaI to generate a single linearised DNA molecule. Lane 3: Marker VI (Roche). Lane 4 Marker II (Roche).



### 3.3.2.1 *in vitro* Cre recombination

Following testing for the functionality of the generated kanamycin resistance cassettes the recombination competence of the mutant loxP sites introduced into one of the cassettes was tested using an *in vitro* Cre recombination reaction followed by a diagnostic restriction digest. 5µg of the pBSLoxNeo prep were incubated with 10µl of Cre-maltose binding protein fusion protein preparation (supplied by Sharon Sheahan, 20mg/ml) in Cre reaction buffer (see appendix 1) in a total volume of 30µl (Kolb and Siddell, 1997). The reaction was incubated at 37°C for 1.5 hours and the DNA subsequently phenol-chloroform extracted and resuspended in 20µl of TE buffer. A fraction of the recombined DNA was set aside for gel electrophoresis without digestion and the remaining fraction was digested with EagI. This restriction enzyme cleaves the kanamycin resistance cassette, generating a 1.9kb linear fragment which will not appear if Cre recombination is unsuccessful. Restriction enzyme digested and uncut plasmid DNA fractions were gel electrophoresed for analysis and banding patterns consistent with successful *in vitro* Cre recombination were observed (figure 3.13) indicating that the introduced mutant loxP sites were capable of interacting.

**Figure 3.13: *In vitro* Cre recombination**



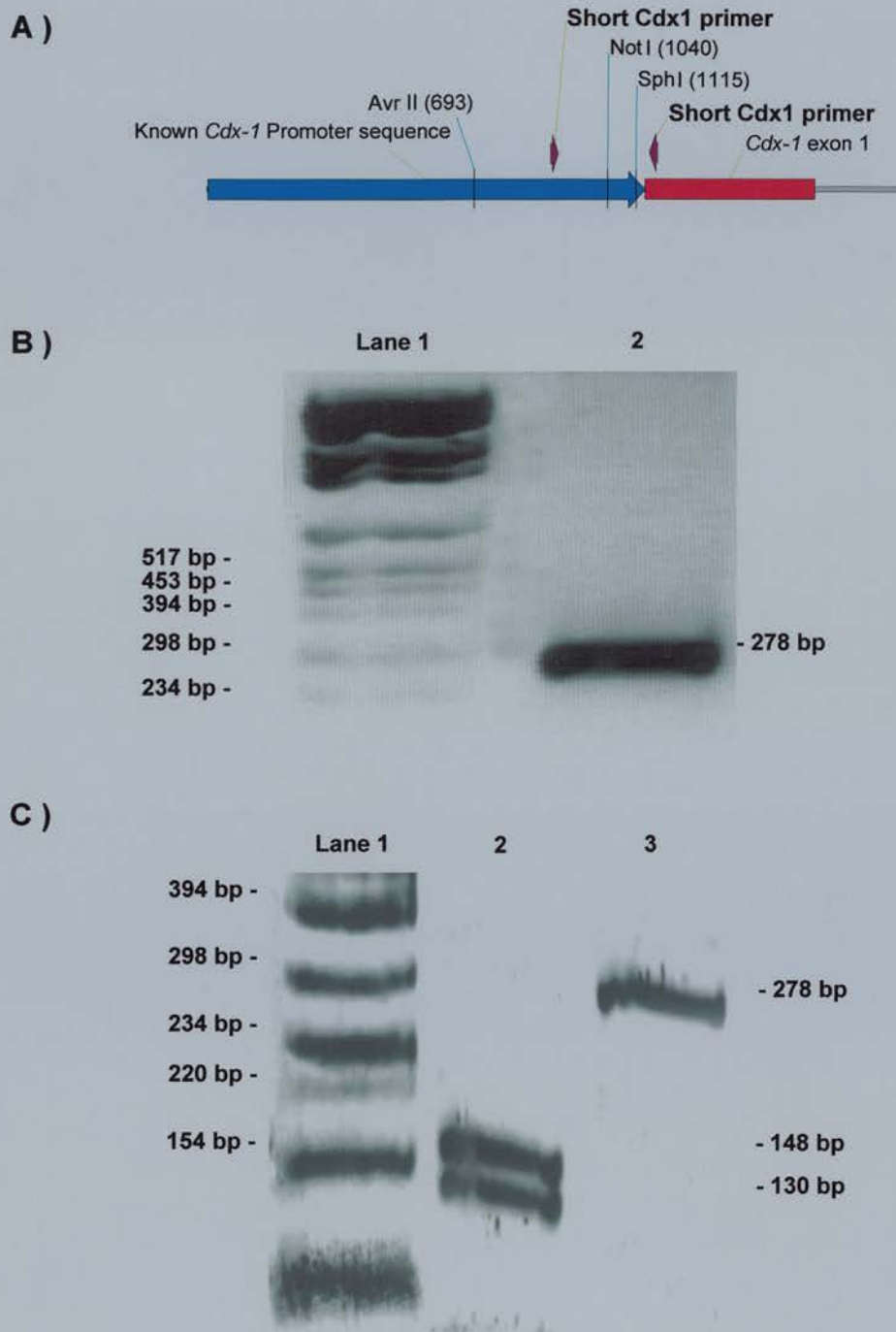
**Figure 3.13: A)** A vector schematic showing the LoxNeo resistance cassette cloned into the pBlueScript plasmid backbone, before and after recombination following *in vitro* Cre treatment. **B)** A gel image showing an analytic EagI digest of the correct pBSLoxNeo plasmid with and without Cre treatment. Lane 1: Marker II (Roche). Lane 2: pBSLoxNeo digested with EagI following Cre treatment. The presence of a band at 1.9 kb indicates that some fraction of the original plasmid has undergone the expected recombination reaction prior to digestion. Lane 3: pBSLoxNeo digested with EagI without prior Cre treatment. Note the absence of the 1.9 kb band. Lane 4: pBS digested with EagI giving a single 2.9 kb band representing the linearised plasmid. Lane 5: The LoxNeo PCR product of 1.9 kb. Lane 6: Marker VI (Roche).

### 3.3.3 PCR and cloning of arms of homology

The final components of a construct for recombination alteration of BAC 219 are the stretches of sequence homologous to the region of the BAC to be replaced: the arms of homology. To generate an appropriate set of homologous sequences for Red $\alpha$ - $\beta$ - $\gamma$  mediated recombination, a PCR was performed across a region of the published *Cdx-1* sequence containing the transcription start signal (figure 3.14, chapter 2 for PCR conditions and primers). The PCR amplified Short-*Cdx-1* sequence (278bp) was cloned into the pCR 2.1 vector (Invitrogen), giving pCRSCdx (see figure 3.15), and sequenced by DNAShef (Royal Infirmary, Edinburgh, UK) to determine the fidelity of the PCR reaction. The sequence was determined to be identical to the corresponding published sequence of murine *Cdx-1* and therefore suitable for use in the generation of arms of homology for Red $\alpha$ - $\beta$ - $\gamma$  mediated BAC recombination. To facilitate further cloning and screening steps, the Short-*Cdx-1* sequence was cut out of the pCR 2.1 vector using a SacI/EcoRV double digest and cloned into the corresponding sites of pBlueScript (Stratagene) giving pBSSCdx (figure 3.16). The amplified Short-*Cdx-1* sequence contains a unique SphI site, at position +73 relative to the transcription start site and 22bp upstream of the translation start site, which was used to introduce the CreER<sup>tm</sup> and *Neo* coding sequences into the Short-*Cdx-1* sequence.

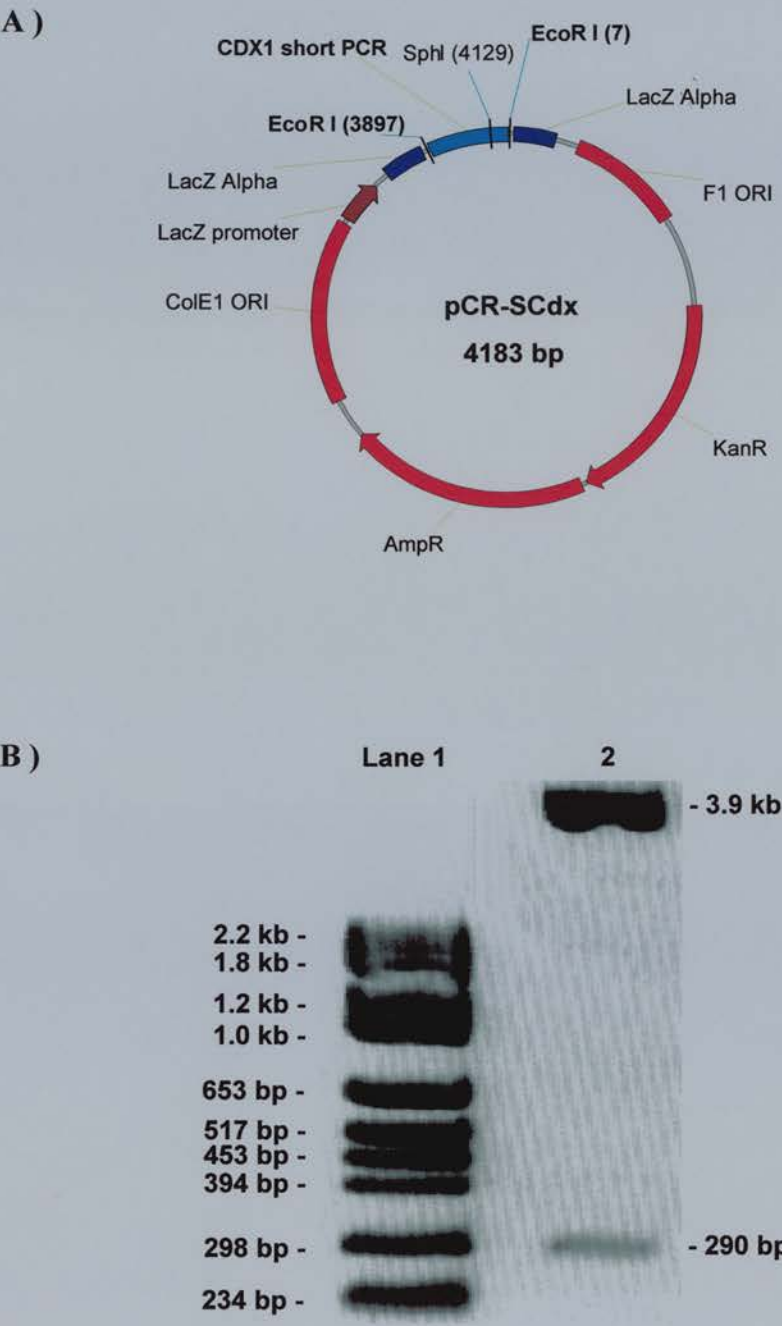


**Figure 3.14: Short *Cdx-1* PCR reaction**



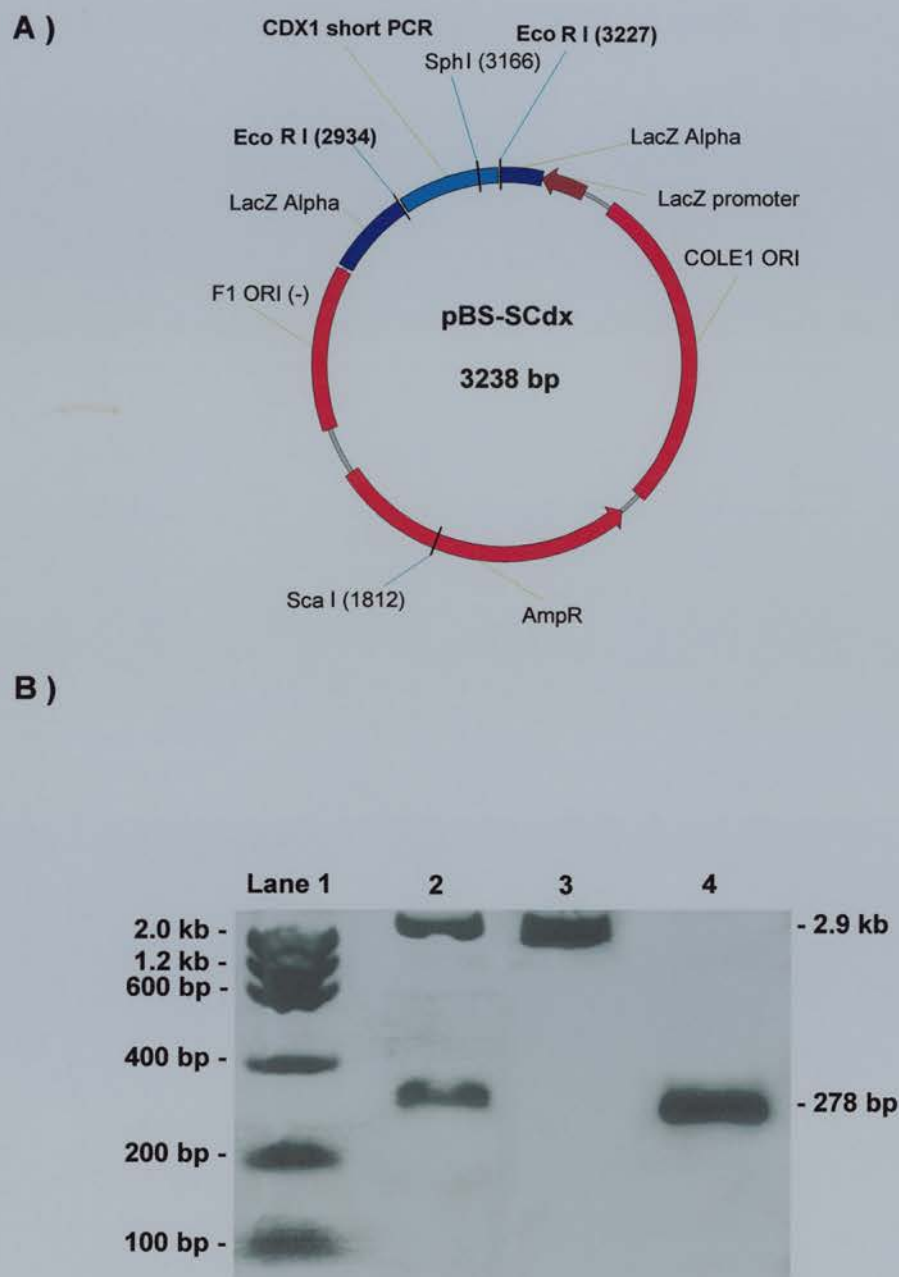
**Figure 3.14: A)** A schematic representation of the region of published murine *Cdx-1* promoter to be amplified as Short-*Cdx1*, showing the location of the PCR primers and some of the restriction sites within the region. **B)** Electrophoretic gel image of an example of the amplified product Short-*Cdx1*, 278 bp in length. Lane I: Marker VI (Roche). Lane 2: Short-*Cdx1* PCR product. **C)** A gel image of an analytic digest of the amplified Short-*Cdx1* product. Lane 1: Marker VI (Roche). Lane 2: Digestion of the correct PCR product with NotI results in the generation of fragments of 148 bp and 130 bp in size. Lane 3: Undigested Short-*Cdx1* PCR product.

**Figure 3.15: Cloning of Short *Cdx-1* PCR**



**Figure 3.15: Cloning of Short *Cdx-1* PCR** **A)** A schematic vector map of the Short-Cdx1 PCR product cloned into the pCR 2.1 vector backbone. **B)** A restriction digest showing the successful cloning of Short-Cdx1 into pCR 2.1 Lane 1: Marker VI (Roche). Lane 2: An analytic *EcoRI* digest of the cloned Short-Cdx1 fragment within pCR 2.1. The *EcoRI* digest liberates the 278 bp Short-Cdx1 DNA sequence with 12 basepairs of vector from the 3.9 kb plasmid backbone.

**Figure 3.16: Subcloning of Short *Cdx-1* PCR**



**Figure 3.16: Subcloning of Short *Cdx-1* PCR** **A)** A vector map of the Short-Cdx1 PCR product subcloned into the pBS vector backbone. **B)** An analytic restriction digest showing the successful subcloning of the Short-Cdx1 sequence into pBlueScript. Lane 1: Low Mass Marker (Roche). Lane 2: An EcoRI digestion of the correct plasmid. The EcoRI digest liberates the 300 bp subcloned Short-Cdx1 fragment from the 2.9 kb pBS plasmid backbone. Lane 3: pBS linearised by EcoRI digestion for size comparison. Lane 3: The 278 bp Short-Cdx1 PCR product. The subcloned fragment is slightly larger as it carries short flanking sequences from pCR 2.1.

### 3.3.4 Construction and screening of final BAC targeting vectors

The three component fragments; the CreER<sup>tm</sup> coding sequence, the floxed or native selection cassette and the Short-*Cdx-1* arms of homology were then assembled into the complete BAC targeting vector. The pBSCre construct was cut with SpeI and treated with Shrimp Alkaline Phosphatase (SAP). The pBSNeo and pBSLoxNeo constructs were digested with XbaI / SpeI (which produce compatible overhang ends) to liberate the resistance cassettes which were gel extracted (see figure 3.17). The pBSCre SpeI linearised fragment and resistance cassette XbaI / SpeI fragments were ligated and the resulting transformed colonies were screened by NotI / BglII digest for the presence of pBSCreNeo and pBSCreLoxNeo plasmids of the correct orientation (see figures 3.18). With the *Neo* resistance cassette inserted in the appropriate direction the entire CreER<sup>tm</sup>-Neo coding sequence could be liberated from the pBS backbone with a NotI digest, giving a band of 4902 bp (with loxP sites) or 4834 bp (without loxP sites) (figure 3.19). The NotI digest was supplemented with T4 DNA polymerase to generate blunt ended fragments and the digest run out for gel extraction of the CreER<sup>tm</sup>-Neo DNA fragment.

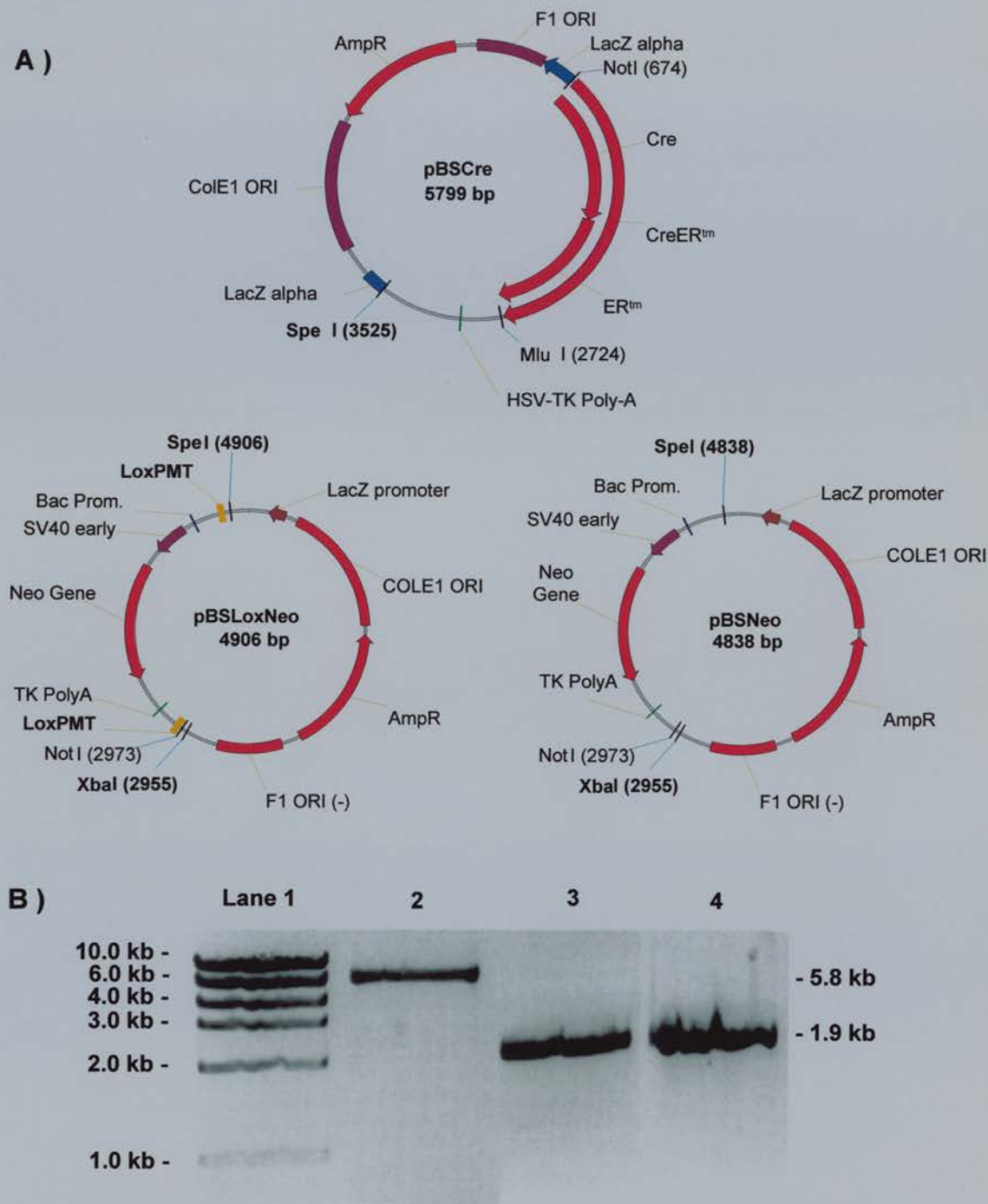
The arms of homology were prepared for insertion of the Cre containing cassette by SphI digest of the pBSSCdx construct, supplemented with T4 DNA polymerase and SAP to generate a blunt ended molecule incapable of self-ligation. This was then ligated to the blunted NotI pBSCreNeo and pBSCreLoxNeo fragments. The resulting plasmids were transformed into TOP10 *E. coli* cells (Invitrogen) and colonies resistant to both kanamycin and ampicillin (the resistance gene present in the pBS backbone) were picked and grown for screening. Plasmids derived from these resistant bacteria were digested with Not I and Bgl II to check for the presence of the expected CreER<sup>tm</sup>-Neo sequence flanked by *Cdx-1* promoter arms of homology and to determine the orientation of the blunted NotI CreER<sup>tm</sup>-Neo fragment within the plasmid (figure 3.20).

Once the appropriate homology arm flanked CreER<sup>tm</sup>-Neo plasmids (pBACtargetLox and pBACtarget) had been identified and isolated they were digested with ScaI, which cut these constructs once to give linearised vectors.

This linearization increases the frequency of recombination in the BAC host as well as preventing intact circular plasmid from persisting within BAC host cells and becoming a source of false positives upon antibiotic selection for resistant recombinants. The linearised plasmids were gel extracted to remove any uncut DNA (figure 3.21). The clean linearised plasmids were then electroporated into previously prepared recombination competent (Red $\alpha$  induced) BAC 219 cells (section 2.1.7).



**Figure 3.17: CreER<sup>tm</sup> and Neo cassettes**



**Figure 3.17: A)** Vector schematics of the previously described pBSCre, pBSLoxNeo and pBSNeo resistance cassette plasmids. **B)** Gel image of the linearised pBSCre and the resistance cassette fragments to be cloned together to generate the CreER<sup>tm</sup>/resistance cassette plasmids. Lane 1: High Mass Marker (Roche). Lane 2: An SpeI digest of the pBSCre which results in its linearisation, giving a single band on 5.8 kb. Lane 3: A gel extraction of the 1.9 kb SpeI/XbaI fragment of pBSLoxNeo plasmid which contains the floxed resistance cassette. Lane 4: A gel extraction of the 1.9 kb SpeI/XbaI fragment of pBSNeo plasmid which contains the non-floxed resistance cassette.

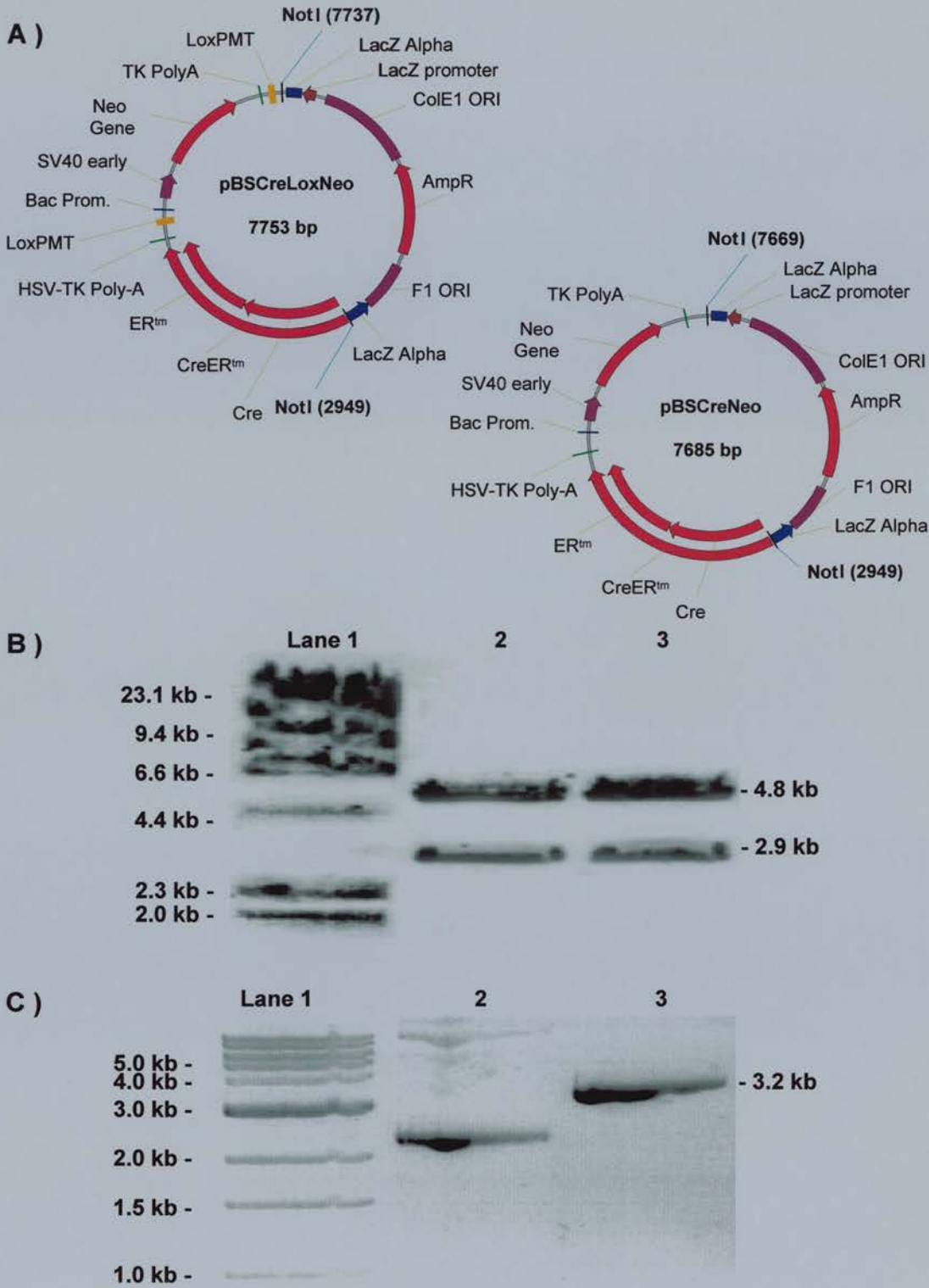
**Figure 3.18: Cre/Neo constructs**



**Figure 3.18: A)** Vector schematics showing the CreER<sup>tm</sup> coding sequence and the Neo resistance cassettes (floxed and native) cloned into the pBlueScript plasmid backbone. **B)** An electrophoretic gel showing the banding pattern of analytic digest of the pBS CreNeo and pBSCreLoxNeo plasmid. Lane 1: Marker II (Roche). Lane 2: pBSCreNeo digested with NotI/BglIII showing the expected bands at 1.5, 3.0 and 3.2 kb. Lane 3: pBSCreLoxNeo digested with NotI/BglIII showing the expected bands at 1.5, 3.0 and 3.25 kb. With the *Neo* cassette inserted in the wrong orientation NotI / BglIII digest would give bands of 4816(floxed)/4884(unfloxed), 1517 and 1352 basepairs.



**Figure 3.19: Isolation of Cre/Neo coding sequence from plasmid backbone**



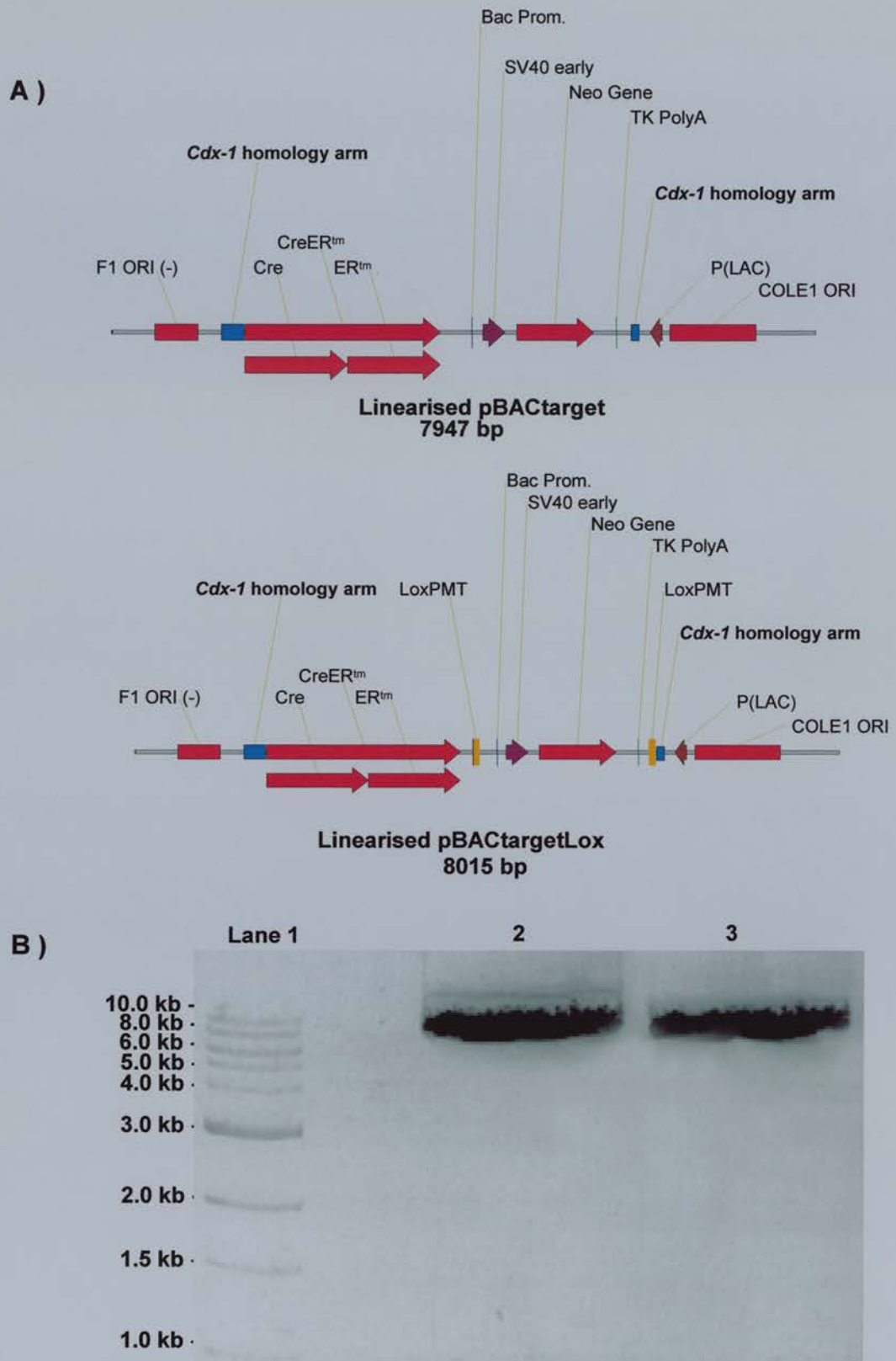
**Figure 3.19: A)** A vector schematic showing the pBSCreLoxNeo and pBSCreNeo plasmids. **B)** An electrophoretic gel showing the banding pattern of a NotI digest of the pBSCreNeo and pBSCreLoxNeo plasmids. Lane 1: Marker II (Roche). Lane 2: pBSCreNeo digested with NotI showing the expected bands at 2.9 and 4.8 kb. Lane 3: pBSCreLoxNeo digested with NotI showing the expected bands at 2.9 and 4.8 kb. **C)** Preparation of the pBS-SCdx plasmid. Lane 1: Kilobase marker (NEB). Lane 2: Uncut pBS-SCdx plasmid. Lane 3: pBS-SCdx plasmid cut with SphI resulting in a linear molecule. T4 DNA polymerase was used to blunt the vector ends in preparation for this cloning step.

**Figure 3.20: BAC targeting vectors**



**Figure 3.20: A)** A vector schematic showing the pBACtarget” and pBACtargetLox plasmids. **B)** An electrophoresis of analytic digests of the blunt cloned plasmids plasmids. Lane 1: Lane 1: Kilobase marker (NEB). Lane 2: pBACtarget digested with BglII and NotI. With the blunted NotI cassette inserted in the correct orientation the expected bands are of 1584 and 6363 bp. Lane 3: A BglII/NotI double digest of the pBACtarget construct with the blunted NotI insert in the wrong orientation; in this plasmid the expected bands are of 3261 and 4686 bp. Lane 4: A BglII/NotI digest of the pBACtargetLox plasmid. The expected bands are of 1584 bp and 6431 bp.

**Figure 3.21: Linearised BAC targeting vectors**

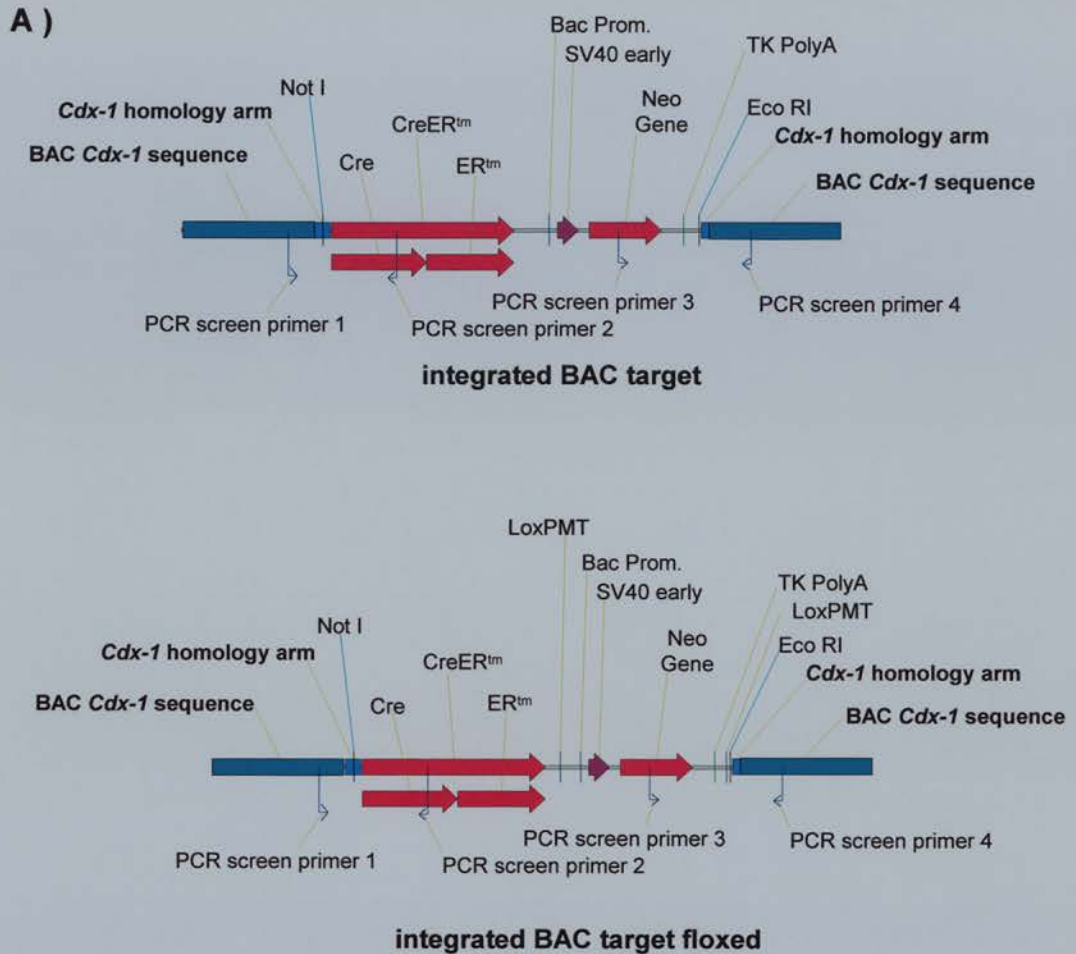


**Figure 3.21: A)** A vector schematic showing the pBACtarget and pBACtargetLox plasmids which have been linearised by digestion with *ScaI*. **B)** An electrophoresis of the *ScaI* digested plasmids. Lane 1: Kilobase Marker (New England Biolabs). Lane 2: *ScaI* digest of pBACtarget. Lane 3: *ScaI* digest of pBACtargetLox.



Initial electroporations of BAC 219/Red $\alpha$  cells with linearised pBACtarget and pBACtargetLox failed to produce colonies on LB plates supplemented with 20 $\mu$ g/ml kanamycin. The incorporation of the *Neo* gene into the BAC would result in only one copy of the resistance gene persisting within the cell, as opposed to the multiple copies present inside host bacteria when the gene is present in a traditional high copy number plasmid, such as the pBS plasmid used as a backbone for the targeting constructs. This marked reduction in copy number may result in a lower threshold of resistance to kanamycin; to compensate, the cells from further electroporations were plated on LB plates containing reduced concentrations of kanamycin. Electroporated and mock electroporated (electroporation without DNA) BAC host bacteria were spread on LB plates containing 1, 2, 5, and 10 $\mu$ g/ml kanamycin. Colonies were observed on 1 $\mu$ g/ml and a reduced number on 2 $\mu$ g/ml plates in both the positive and negative electroporations. No bacteria grew on plates supplemented with 5 or 10 $\mu$ g/ml kanamycin, regardless of the presence of either linearised BAC targeting construct in the electroporation. Further experiments showed that a very small number of colonies (1-5) grew on plates supplemented with 3 $\mu$ g/ml kanamycin, again apparently independent of the presence of the targeting construct in the electroporation. The small number of colonies found on these 3 $\mu$ g/ml kanamycin plates, both positive and negative, were screened for the presence of the recombination modified BAC. As the isolation of BAC construct in quantities suitable for analysis by restriction digest is a relatively time consuming process, the possible targeted BAC clones were screened by PCR with primers located in the sequence to be inserted into the BAC and within the *Cdx-1* sequence previously confirmed to be contained within BAC 219 in such a manner that the product from the unaltered BAC and that generated from the targeted BAC are of different sizes, 950 and 1550 bp respectively (figure 3.22, section 2.2.2 for PCR primers and conditions). All colonies isolated from these initial electroporation experiments proved to contain the unmodified BAC 219.

**Figure 3.22: Integrated BAC transgene constructs**



**Figure 3.22: A)** A vector schematic showing the expected products after Red $\alpha/\beta$  mediated recombination of the “BAC target” and “BAC target floxed” sequences into BAC 219.

The lack of expected recombination products from the electroporations could have had several causes. The first possibility investigated was the level of electroporation competence of the BAC 219 host cells. These cells were made electrocompetent by a commonly used osmotic shock protocol (see section 2.1.7). This protocol had proved relatively inefficient ( $\sim 5 \times 10^3$  colonies/ $\mu\text{g}$  of plasmid DNA) for preparing BAC 219 hosts for the introduction of the inducible pBAD-Red $\alpha\beta\gamma$  recombination plasmid; the additional requirement of a recombination event in the attempt to create modified BACs would lower the number of resistant colonies generated in an electroporation with a given amount of DNA. The osmotic shock protocol was altered in an effort to improve the electroporation competence of the BAC 219 hosts; additional resuspension steps (see section 2.1.7) with both dH<sub>2</sub>O and 10% glycerol improved the electroporation efficiency slightly ( $\sim 1 \times 10^4$  colonies/ $\mu\text{g}$  of pBS DNA), as did the use of 10mM CaCl<sub>2</sub> in place of dH<sub>2</sub>O for one of the resuspension steps ( $\sim 1.25 \times 10^4$  colonies/ $\mu\text{g}$  of pBS DNA), although these additional steps reduced the amount of bacterial prep which could be recovered and aliquoted. Electroporations using linearised pBACtarget and pBACtargetlox with these more electrocompetent bacteria produced slightly higher numbers of colonies resistant to 3 $\mu\text{g}/\text{ml}$  kanamycin, however all of these colonies proved to contain unmodified BAC 219 when screened by PCR. Another possible cause of the inability to generate recombination altered BAC constructs could lie in the *Neo* resistance cassette used in the pBACtarget vectors. This resistance cassette did, however, successfully generate kanamycin resistance in all host bacteria used in the various cloning stages involved in construction of the final pBACtarget and pBACtargetlox vectors, suggesting that the cassette was functional and being expressed at sufficient levels to provide resistance to 20 $\mu\text{g}/\text{ml}$  kanamycin when present within a cell on a high copy number plasmid.

Variation of the electroporation parameters (see section 2.1.7) failed to increase the efficiency of the transformation. Lowering of the electroporation voltage to 2.3kV from the recommended 2.5kV, the maximum potential generated with the electroporation apparatus available, resulted in a small decrease in the



transformation efficiency, while variation in the capacitance and/or resistivity of the system resulted in very large reductions in efficiency.

An additional consideration was the possibility that the Red $\alpha$  protein was not being induced to sufficient levels in the BAC 219 hosts prior to the osmotic shock treatment. To check for induction of the Red $\alpha$  protein, 1.0 ml aliquots of the BAC 219 host culture were taken at 30 minute intervals following the addition of 0.2% L-(+)-arabinose (w/v) to the culture medium. Protein preparations were made by the pelleting of bacterial cells at 10,000g for 2 minutes in a benchtop microcentrifuge, followed by resuspension in 100 $\mu$ l of SDS-protein loading buffer. The bacterial cells were lysed by placing the samples in a boiling water bath for 5 minutes after which the protein samples and appropriate molecular weight markers were electrophoresed on a 10% acrylamide gel. The separated proteins were stained with GelCode (Peirce). A clear increase of the intensity of a band corresponding to the apparent molecular weight of Red $\alpha$  (25kDa) was seen in protein samples collected as little as 1 hour and continuing at least 2 hours following the addition of arabinose to the bacterial culture when compared to a sample collected immediately prior to the addition of the pBAD inducer (see figure 3.5). This indicates that production of the Red $\alpha$  protein was being successfully induced in the BAC 219 host cells well within the 2 hours between addition of arabinose to the culture medium and harvesting of the cells for osmotic shock treatment. While this would appear to indicate that the Red $\alpha$  protein was present in significant quantities in the BAC host bacteria, this gives no indication of the quantities or indeed the presence of the Red $\beta$  or Red $\gamma$  proteins. The genes encoding these two viral proteins were under the control of constitutive prokaryotic promoters, EM7 and Tn5 respectively, both strong promoters (Goldstein and Doi, 1995). It is possible that the levels of transcription from these promoters, and hence the concentrations of the Red $\beta$  and Red $\gamma$  proteins, were too low to achieve a significant rate of recombination in arabinose activated BAC host cells electroporated with the BAC targeting vectors. Alternatively, mutations in the constitutive promoters, or in any of the protein coding sequences, could have resulted in a lack of

functional recombination promoting proteins. While these possibilities seem unlikely, they could not be ruled out.

A further consideration for electroporation efficiency is the concentration and quality of the DNA used in the electroporation. The *ScaI* linearised pBACtarget and pBACtargetlox plasmids were gel extracted using the QIAEX II gel extraction kit (Qiagen) in order to prevent electroporations from being contaminated with uncut plasmid. The gel extraction of the 8kb DNA molecules resulted in relatively low yield (~50%), the DNA concentrations of the extracted preparations were typically between 0.25 and 0.4µg/µl despite efforts to improve yield. This concentration is relatively low but within the limits for electroporation experiments. To determine if increased DNA concentration would produce better electroporations, pBACtarget and pBACtargetlox were digested overnight with a large excess (20 units/µg) of *ScaI* to reduce the amount of uncut plasmid and the digests were phenol:chloroform extracted. The pelleted DNA was resuspended in a minimal volume (10µl) of 10mM Tris (pH 8.0). This resulted in DNA concentrations of 1-2µg/µl; electroporations using this concentrated DNA prep again produced a slight increase in the number of colonies resistant to 3µg/ml kanamycin, although all colonies screened by PCR contained unmodified BAC 219. In order to determine if the increase in colonies resistant to 3µg/ml kanamycin were the result of contamination of the DNA prep with uncut plasmid a selection of the resistant colonies were screened by PCR for the presence of the *Neo* resistance gene. The presence of the *Neo* resistance coding sequence in a number of the colonies which also contained unmodified BAC 219 suggested that these bacteria were indeed taking up contaminating uncut plasmid from the phenol:chloroform DNA prep.

In order to achieve higher concentrations of linear BAC target sequence without the uncut plasmid contamination a PCR reaction was designed to generate a BAC targeting molecule similar to the cloned pBACtarget. Primers were designed which contained 40bp *Cdx-1* arms of homology (5' and 3') in addition to sequences designed to anneal to the ends of the CreER<sup>tm</sup> / *Neo* resistance coding sequence. The PCR reaction would use the pBSCreNeo plasmid as a template

(see figure 3.18), generating high concentrations of linear DNA molecules containing the CreER<sup>tm</sup>-Neo sequence flanked by short arms of homology. A consideration in performing this PCR reaction was the introduction of sequence errors by the high temperature polymerases commonly used in PCR reactions. Mutations in the *Neo* resistance gene would be relatively unimportant, as non-functional *Neo* coding sequence would be selected against on introduction into bacterial hosts. Mutations in the Cre coding sequence would be of greater consequence, however, as these alterations may lower or negate the function of the resulting Cre protein but would not be detected without considerable sequencing or functional analysis. In an effort to reduce the possibility of polymerase error, the proofreading high temperature polymerase Pfu was used in this PCR reaction. A number of reaction conditions were used in an attempt to generate the appropriate BAC targeting molecule, however no product was detected. The difficulty in amplifying this product may have been due to the high secondary structure present in the upstream primer sequence; the translation start site of *Cdx-1*, the most appropriate location to attempt to place the translation start site of the CreER<sup>tm</sup>-Neo sequence, is found just downstream of a GC repeat island (see appendix 2). This DNA motif, common to many eukaryotic promoter sequences, meant that the primer designed to overlap this area could self-anneal, as well as being able to anneal to other copies of the same primer. A number of variations in the annealing temperature of the PCR reaction sequence were used in an effort to overcome the effects of this poly-GC sequence but no annealing temperature resulted in detectable levels of product.

A further issue is the linearisation of the pBACtarget(lox) plasmids prior to electroporation into the activated BAC host bacteria. The rate of homologous recombination promoted by a bacterial system analogous to the phage system used in this experiment is known to be increased by the proximity of double strand breaks to the region of homology (Keim and Lark, 1990). Using PCR to generate a BAC targeting construct leaves DNA strand ends very near the arms of homology; the plasmid vectors used here were linearised using a restriction digest which leaves substantial (1-2kb) stretches of DNA between the restriction generated strand ends and the arms of homology. This may have significantly reduced the efficiency of the recombination reaction; unfortunately the nature of

the pBACtarget(lox) vectors prevented further digests to remove this “extra” sequence prior to electroporations.

### 3.4 Discussion

The inability to detect the expected recombination products from a large number of electroporations highlights some of the potential difficulties arising from using this system. There are a number of factors which should be taken into consideration before attempting to generate altered BAC construct by recombination. The first is the selection of an appropriate antibiotic resistance cassette. The BAC host bacteria used in this series of experiments have a small degree of natural resistance to kanamycin which only became apparent when the concentrations of the antibiotic were lowered to compensate for the single copy of the *Neo* resistance gene expected to be present within the targeted BAC. This small natural resistance resulted in relatively large numbers of untransformed colonies growing on plates intended to select for recombinants; colonies resulting from the relatively rare recombination events may have been “hidden” amongst these unrecombined colonies. It proved impossible to find an antibiotic concentration which would return colonies from electroporations involving DNA while fully inhibiting growth in mock-electroporated controls. It is possible that other antibiotic resistance cassettes would have more fully inhibited growth of non-recombinants at the low concentrations needed to allow growth of cells containing BACs bearing a single copy of the cassette.

Another consideration for future efforts to create BAC based constructs through recombination is the ability to generate concentrated, high quality linearised targeting DNA preparations. PCR reactions can generate the required concentrations and DpnI (a methylation dependant restriction enzyme) digestion of the amplification reaction can be used to eliminate the methylated template plasmid DNA. However PCR reactions may introduce changes to the sequence to be recombined into the BAC. This may not be crucially important if the material to be introduced into the BAC consist solely of a selectable marker, but it becomes more important if the sequence contain other genes which are not so easily tested for their ability to generate functional protein product. The use of proofreading polymerases may partially circumvent this problem but may still require time consuming sequencing reactions to verify crucial sequence integrity. Additionally, if the sequence to be inserted into the BAC does not contain an

endogenous promoter and is intended to make use of promoter material already present within the sequence of the BAC, the design of efficient PCR primers may be hindered by the presence of repetitive motifs in some promoter sequences.

The ability to generate BAC based transgenes within their bacterial hosts through induced recombination holds much promise. BAC based transgenes may be large enough to escape the positional effects which can alter the efficiency of expression from smaller, traditional transgenes. The amount of sequence present within BACs, up to 120kb in some cases, may also allow the creation of transgenes based on genes or promoters whose sequence has not been fully described, as was the intention in this effort to generate a transgene based on the promoter of the *Cdx-1* murine homeobox gene. While these features make the building of transgenes using *in vivo* recombination systems and BAC constructs attractive, the nature of the transgenic material and the careful choice of selection markers are vitally important to success.



# **Chapter 4**

## **Evaluation of murine *Cdx-1* promoter sequence**

## Chapter 4 – Evaluation of murine *Cdx-1* promoter sequence

### 4.1 Background

The construction of the BAC 219 based transgene discussed in the previous chapter would have allowed for the transcription of the CreER<sup>tm</sup> recombinase to be placed under the control of a large stretch of the native murine *Cdx-1* promoter sequence. It is hoped that this would allow for an accurate reproduction of the time and tissue specificity of the *Cdx-1* expression patterns, thereby imparting an important degree of control over the activity of the Cre recombinase. In parallel with the efforts to alter BAC 219 to include Cre coding sequence, a more traditional transgene construct was prepared by Dr. Sula Corbet and Dr. Jane Armstrong using a subcloned fragment of BAC 219.

The *Cdx-1* promoter has been partially characterised in at least three separate studies, two of these studies have been published since this project was begun. In the first, the stretch of sequence 1040bp in length immediately upstream of the transcription start site was used to drive expression of a chloramphenicol acetyltransferase (CAT) based reporter construct in a number of human derived colorectal carcinoma cell lines as well as murine NIH 3T3 cells (Hu *et al.*, 1993). Using a series of deletion constructs as well as subcloned stretches of the 1040bp and sequence analysis, a number of transcriptionally important motifs were identified; these include a silencer element found between –589 and –380, a silencer blocking element between –1040 and –887 and a positive control element found between –47 and +66, all positions relative to the transcription start site (Hu *et al.*, 1993). This last positive element was the only one of the identified elements to confer cell type specificity; the element preferentially drove transcription in epithelial cell lines (see figure 4.1) (Hu *et al.*, 1993). While information regarding the functionality of this short stretch of promoter was gained, the levels of expression given by this short stretch of sequence were extremely low, suggesting that more promoter was necessary to drive adequate transcription or that the cell lines used in the experiment were capable of down regulating the activity of the *Cdx-1* promoter; downregulation of *Cdx-1* transcription has been observed in a number of epithelial derived human cell lines (Mallo *et al.*, 1997).

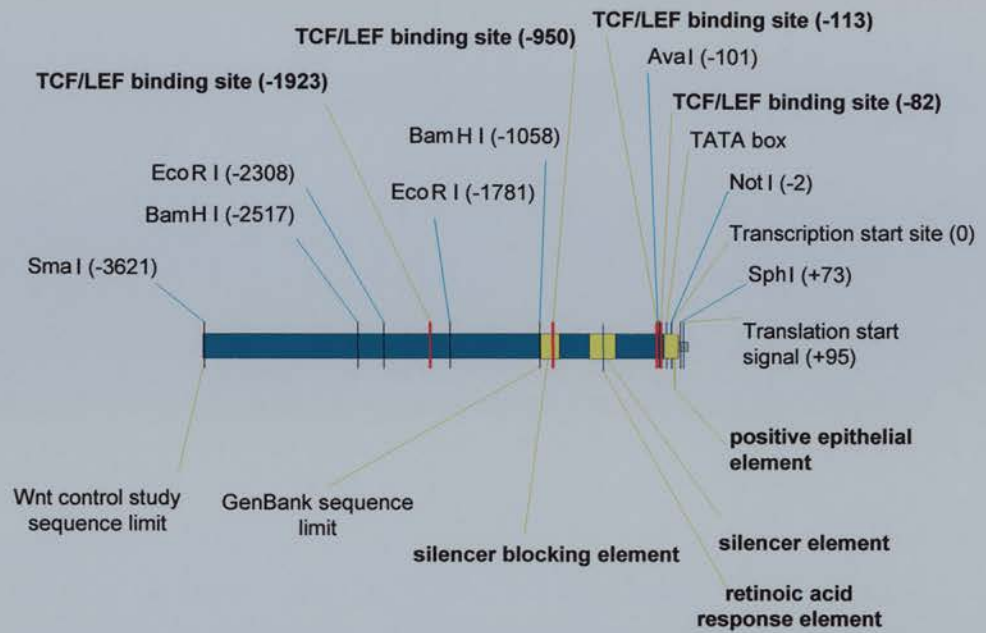
In the second study a larger stretch of sequence, comprising some 3621bp upstream of the transcription start site, was used in the construction of a number of luciferase based reporter plasmids which were transfected into wild type and Tcf-4 deficient murine ES cells as well as undifferentiated rat embryonic endoderm (Lickert *et al.*, 2000). An analysis of the sequence of this stretch of DNA had revealed, in addition to the elements described in the earlier study, four perfect TCF/LEF binding sites (base pair motif CTTGA(T)A(T)) (see figure 4.1). The presence of these sequence motifs suggested that the expression of *Cdx-1* was on some level under the control of the Wnt signalling pathway. In this study the experiments focussed on the role of these transcription-factor binding sites in the control of transcription from the *Cdx-1* promoter. Each of the TCF/LEF sites were in turn removed and/or mutated to abrogate Tcf binding. Changes in the levels of reporter protein generated from the various plasmids revealed the contributions of each of the TCF/LEF sites to overall transcriptional level alterations induced by Tcf binding and transactivation. The results suggested that the two TCF/LEF binding motifs nearest the transcription start site at positions -113 and -82 contributed most strongly to the transactivation effects of Wnt signalling, while the two more distant elements at positions -1923 and -950 contributed considerably less to Wnt activation of *Cdx-1* transcription (Lickert *et al.*, 2000).

The third and most recent investigation reveals another transcription factor influencing expression driven by the *Cdx-1* promoter sequence (Houle *et al.*, 2000). Similarities between the altered vertebral specifications seen in *Cdx-1* mutant mice and the phenotype of mice mutant for retinoic acid response elements (RAREs) suggested a possible overlap in the functions of homeobox genes and retinoic acid responsive genes. While no consensus RARE was found in the sequence of the *Cdx-1* promoter, application of retinoic acid to a number of cell types resulted in specific upregulation of *Cdx-1* levels. The 15bp (AAGGGTCGTGACCCT) element within the *Cdx-1* promoter responsible for the binding of retinoic acid was found at -512 relative to the transcription start site. This lies within the silencer element identified in the initial isolation and characterisation of the *Cdx-1* mRNA and genomic promoter. Interestingly another study has shown that  $\beta$ -catenin signalling through the TCF/LEF pathway

can be inhibited by the presence of retinoic acid and that  $\beta$ -catenin is capable of binding the retinoic acid receptor RAR (Easwaran *et al.*, 1999a). Because *Cdx-1* regulates homeobox genes which themselves are not responsive to concentrations of retinoic acid, the suggestion is that *Cdx-1* acts as a transducer of the retinoic acid signals to influence vertebral patterning through regulation of these homeobox genes.

In the present study a large stretch of the DNA immediately upstream of the transcription start site, representing some 6.7kb of sequence, was analysed for relevant promoter elements while reporter constructs based on the Enhanced Green Fluorescent Protein (EGFP) were constructed using DNA fragments of up to 4.3kb in length. These stretches of promoter DNA were all derived from the BAC clone 219.

**Figure 4.1: published *Cdx-1* promoter elements**



**Figure 4.1:** Schematic of the previously described regulatory elements of the *Cdx1* promoter. The investigated elements include 4 TCF/LEF binding sites, a novel retinoic acid response element, a orientation independent silencer and its blocking element and an element giving epithelial cell type transcription specificity. The 5' end of the diagram indicates the limit of previously characterised sequence.

## 4.2 Mapping of newly isolated upstream *Cdx-1* promoter sequence

The first step to isolating novel sequence upstream of the *Cdx-1* gene was to construct a map of restriction sites within BACs 219 and 538, both of which were determined to contain *Cdx-1* sequence by PCR and hybridisation (see section 3.2.1). To this end Drs. Sula Corbet and Jane Armstrong digested BAC 219 and 538 DNA preparations with a number of single or paired restriction enzyme reactions. The resulting fragments of BACs 219 and 538 were then run out and visualized on an ethidium laced agarose gel (see figure 4.2a). The pattern of the digestion products allowed a rough map of restriction sites to be generated and similarities confirmed that the two BACs were overlapping clones, as would be expected if both contained *Cdx-1* exonic sequence as previously determined. The BACs both appeared to contain at least 100kb of sequence; in order to determine which of the bands produced in the various restriction digests contained the desired sequence the gel was blotted and probed with the DNA sequence of *Cdx-1* exon 1 by Dr. Sula Corbet. A number of bands hybridised to this probe (see figure 4.2a, red) indicating that they contained at least the sequence of the first exon of *Cdx-1*. The restriction map of the known promoter sequence and the first exon allowed for the determination of which bands would prove to contain a relatively large amount of more upstream sequence. It was ultimately decided that a band generated from the BAC 219 XhoI digestion, approximately 7.1kb in size, provided the best target for subcloning. The band proved reasonably separated from other bands resulting from the same digest allowing for gel extraction and was confirmed by PCR to contain the first exon of *Cdx-1* (see figure 4.2b). The genomic sequence downstream of the *Cdx-1* promoter contains an XhoI site at position +640 relative to the transcription start site; as a result the band of 7.1kb produced by the XhoI digest would contain approximately 5.8kb of sequence upstream of the -1040 position, which was the limit of sequence data at the start of this project (GenBank). The 7.1kb BAC 219 XhoI band was gel extracted and cloned into the corresponding site of the pGEM commercial plasmid vector (Promega) by Dr. Sula Corbet. Drs. Jane Armstrong and Sula Corbet have since used this subcloned section of BAC 219 in the construction of a CreER<sup>tm</sup> transgene while this project used the pGEM-XhoI sequence to investigate the properties of the *Cdx-1* promoter.

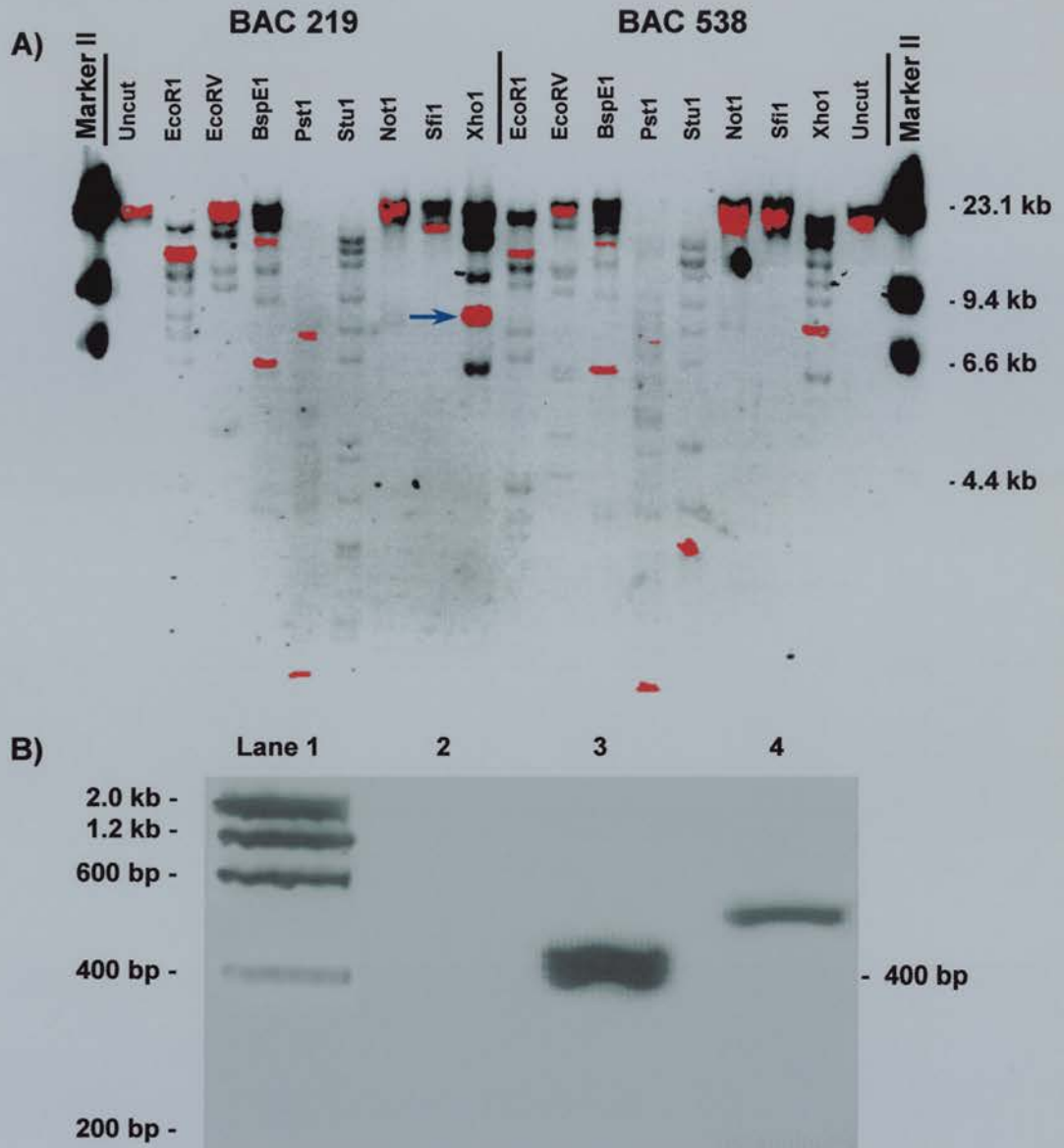


### 4.3 Subcloning and sequencing of pGEM-XhoI

The restriction digestion of the BAC 219 and 538 provided some degree of mapping of the construct but further mapping was necessary in order to use the subcloned fragment of BAC 219 for subsequent investigations, including sequence analysis and reporter construct assembly. The pGEM-XhoI construct was prepared and digested with a number of individual endonucleases as well as in double digestions using XhoI in conjunction with a second restriction endonuclease. A number of the individual digestions were able to provide information on orientation of the fragment within the pGEM backbone, while the double digests with XhoI provided the approximate locations of restriction sites within the subcloned BAC fragment. Further double digests allowed for more accurate placement of sites relative to one another (see figure 4.3a). This more accurate map was then able to form the starting point for subcloning sections of the BAC 219 XhoI fragment in order to efficiently sequence its entire length as well as allowing the simultaneous construction of several EGFP based reporter vectors.

Initial sequencing of the pGEM-XhoI plasmid was done using a primer within the known *Cdx-1* promoter sequence and a primer within the pGEM plasmid backbone sequence at the upstream end of the XhoI fragment (DNAShef, Royal Infirmary, Edinburgh). Further sequencing was done using the same plasmid backbone primer following digestion of the pGEM-XhoI plasmid with SacI and EcoRI and subsequent religations. These digestions removed portions of the upstream XhoI fragment sequence and allowed determination of internal sequence. Subsequent sequencing was performed using primers complementary to this newly determined internal sequence. The DNA sequence data was compared and aligned to reconstruct the full length of the BAC 219 XhoI fragment (see figure 4.3b, appendix 2 for full sequence data).

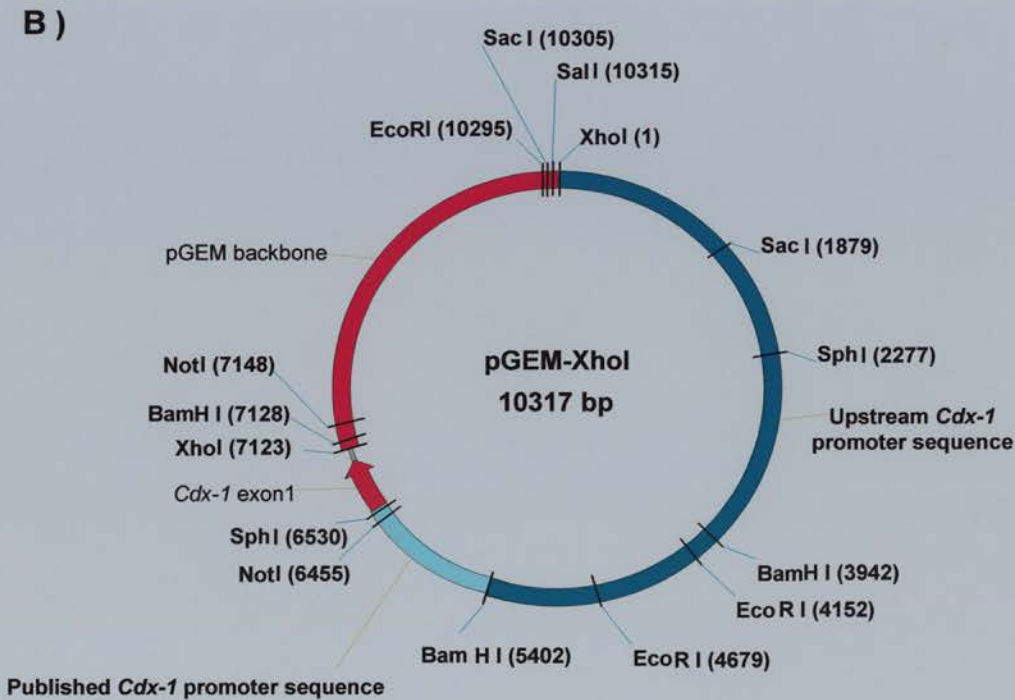
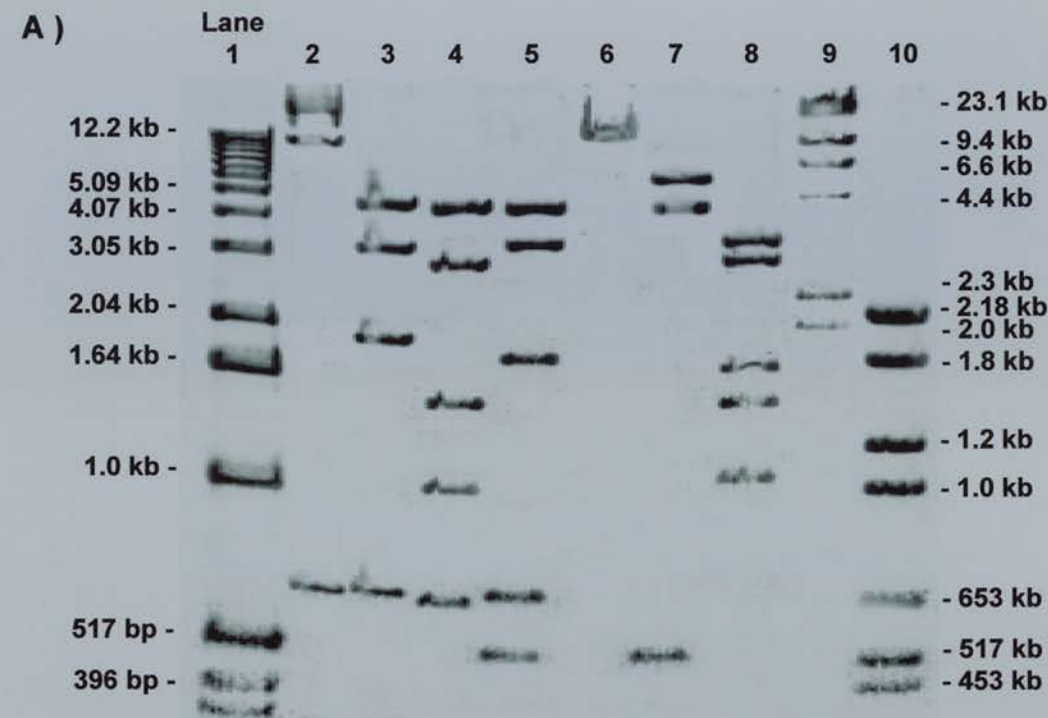
**Figure 4.2: BAC subcloning**



**Figure 4.2: A)** EtBr stained gel showing the banding patterns following digestion of BACs 219 and 538 with the restriction endonucleases indicated above the appropriate lanes. The bands common to both BACs suggest that they contain significant amounts of identical sequence other than that of the construct backbone. Lanes 1 and 20 contain Marker II (Roche). The gel image is overlaid with a false colour representation of the bands observed following southern hybridization of the BAC 219 and 538 digests using *Cdx1* exon 1 as the probe sequence. **B)** PCR confirmation of the 7.1 kb *XhoI* band found in both the BAC 219 and BAC 538 digestions as containing *Cdx-1* exon 1. The *XhoI* bands at ~10.0 kb, 7.1kb and 6.0 kb were gel extracted and used as template for the PCR amplification (see section 2.2.2) of *Cdx-1* exon 1. Only the 7.1 kb band (indicated by the blue arrow) gives the correctly sized band. Lane 1: Low Mass Marker (Gibco). Lane 2: *Cdx-1* exon 1 PCR reaction using gel extraction of 6.0 kb BAC 219 *XhoI* band. Lane 3: *Cdx-1* exon 1 PCR reaction using gel extraction of 7.1 kb BAC 219 *XhoI* band. Lane 4: *Cdx-1* exon 1 PCR reaction using gel extraction of 10.0 kb BAC 219 *XhoI* band.

**Figure 4.3: A)** Gel image of some of the restriction digests of pGEM-XhoI plasmid used to determine the orientation of the insert within the plasmid backbone as well as determining the number and the approximate positions of various sites within the insert itself. Lane 1: Kilobase Marker (Gibco). Lane 2: NotI digest, the presence of the 665 basepair band shows insert orientation within the pGEM plasmid. Lane 3: NotI/SacI digest. This digest liberates the insert from the 3.2 kb pGEM backbone, again shows the NotI 665 basepair fragment and reveals one SacI site within the BAC 219 insert sequence, which cuts the insert into 4.5 and 1.9 kb fragments. Lane 4: NotI/BamHI digest of pGEM-XhoI. The expected band at 645 bp (20 less than the 665 bp NotI fragment) appears, the pGEM backbone and a portion of the insert run at just over 4.0 kb and the digest shows the presence of 2 more BamHI sites within the remaining BAC 219 insert sequence which give bands of ~2.7, 1.5 and 1.0 kb. Lane 5: NotI/EcoRI digest of pGEM-XhoI. This digest again liberates the insert from the 3.2 kb pGEM backbone and shows the NotI 665 basepair fragment. Additionally 2 EcoRI sites are shown to exist within the BAC 219 sequence, cutting the insert into 4.2, 1.8 and 0.5 kb fragments. Lane 6: SalI digest of pGEM-XhoI, a single SalI site exists in the pGEM multiple cloning site and no SalI sites appear within the subcloned insert sequence. Lane 7: SalI/EcoRI digest. The pGEM backbone and some of the insert travel at 5.6 kb while the insert is cut into 4.2 and 0.5 kb fragments. Lane 8: SalI/BamHI digest. This digest liberates the pGEM vector (3.2 kb) and reveals the approximate locations of the BamHI sites within the insert sequence. Lane 9: Marker II (Roche). Lane 10: Marker VI (Roche). **B)** A vector schematic of the BAC 219 7.1 kb XhoI fragment cloned into the pGEM plasmid backbone, showing the positions of some restriction endonuclease sites as determined by the described digests in conjunction with derived sequence data.

**Figure 4.3: Mapping of pGEM-XhoI**



#### 4.4 Reporter plasmid construction

The newly derived upstream *Cdx-1* promoter sequence contained all of the transcriptional mediating elements described in earlier studies (see section 4.1). In addition to the described elements, a fifth perfect TCF/LEF binding site, not reported previously, was found at position -3957 relative to the transcription start site. In order to investigate transcription initiation from the newly isolated *Cdx-1* sequence a number of reporter constructs were assembled for use in *in vitro* studies. A number of reporter systems were considered and compared for use in the *Cdx-1* promoter characterisation experiments. The most suitable reporter system proved to be one based on the EGFP gene. The EGFP gene is a specifically mutated version of the Green Fluorescent Protein (GFP) originally isolated from the jellyfish *Aequorea Victoria* (Cormack *et al.*, 1996; Prasher *et al.*, 1992). The GFP gene encompasses 714bp encoding a protein of 238 amino acids which fluoresces green at 508nm under stimulation with UV light (absorption wavelengths of 395 and 470nm). A number of features of GFP make it useful as a reporter system; the GFP protein appears to be relatively non-toxic in a large number of cell types, it fluoresces without the need for cellular co-factors and GFP fusion proteins (both N- and C- terminal) retain the ability to fluoresce enabling tracking of specific proteins within a cell (Wang and Hazelrigg, 1994). These properties have been retained and added to in the enhanced version of the GFP protein. EGFP remains relatively non-toxic and can form C- and N-terminal fusion proteins without loss of chromophore activity. Its absorption spectrum has been shifted to 488nm by two amino acid changes (Phe64 to leucine and Ser65 to threonine) while the emission spectrum remains similar to that of native GFP, with the major peak at 508nm (Cormack *et al.*, 1996). The transition to Leu64 and Thr65 also boosts the level of fluorescence and prevents the incidence of photobleaching making EGFP a more sensitive marker protein than GFP (Cormack *et al.*, 1996). Further alterations of the original GFP sequence includes silent alteration of a large number of codons to human preference versions and the addition of a Kozak sequence in order to increase expression of the EGFP protein in many mammalian cell types (Kozak, 1987a; Kozak, 1987b). The EGFP coding sequence has been incorporated into a number of commercially available vectors for use in expression and localisation



studies. A vector encoding a promoterless version of EGFP, pEGFP-1 (Clontech), was used as the backbone for construction of the *Cdx-1* reporter constructs and as a negative control for expression studies (see figure 4.4).

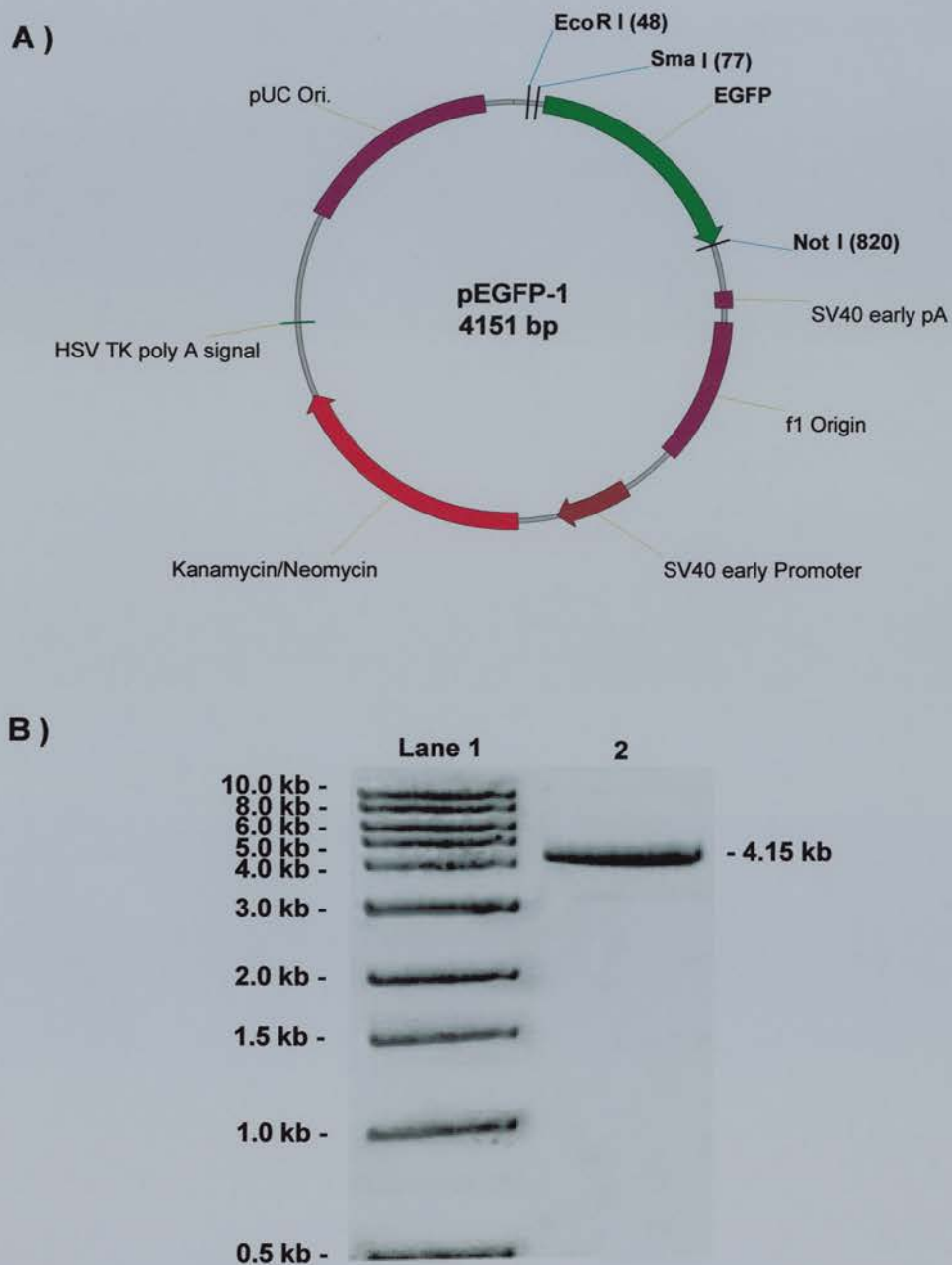
Analysis of the newly isolated *Cdx-1* promoter sequence revealed an SphI site at position -4186 relative to the transcription start site. Another SphI site is found within the published sequence, at position +73 relative to the transcription start site and 22bp upstream of the translation start site. The sequence between these two SphI contains all the published *Cdx-1* promoter elements as well as the newly identified fifth TCF/LEF binding motif, while excluding the translation start site (+95) (see figure 4.5a). The 4.2kb SphI fragment of *Cdx-1* promoter was chosen as the starting point for the construction of the various *Cdx-1*-EGFP reporter plasmids.

#### **4.4.1 Construction of pEGFP-Sph**

Digestion of the pGEM-XhoI plasmid with SphI liberates the 4.2kb *Cdx-1* promoter sequence from the 3.2kb pGEM backbone and 2.5kb of the 5' end of the BAC 219 XhoI fragment (see figure 4.5b). Unfortunately, the pEGFP-1 plasmid does not contain a site within the multiple cloning site upstream of the EGFP coding sequence which generates ends compatible with those generated on SphI digestion. To clone the 4.2kb SphI fragment into the pEGFP-1 vector the digestion was supplemented with dNTPs and T4 DNA polymerase in order to generate a blunt ended molecule. This fragment was cloned into the blunt SmaI site found in the pEGFP-1 multiple cloning site and the resulting plasmids screened for the presence and orientation of the 4.2kb SphI *Cdx-1* promoter fragment (see figure 4.6). The resulting 8.2kb plasmid, pEGFP-Sph, contains the EGFP coding sequence under the control of all 5 TCF/LEF binding motifs as well as all other published promoter elements.

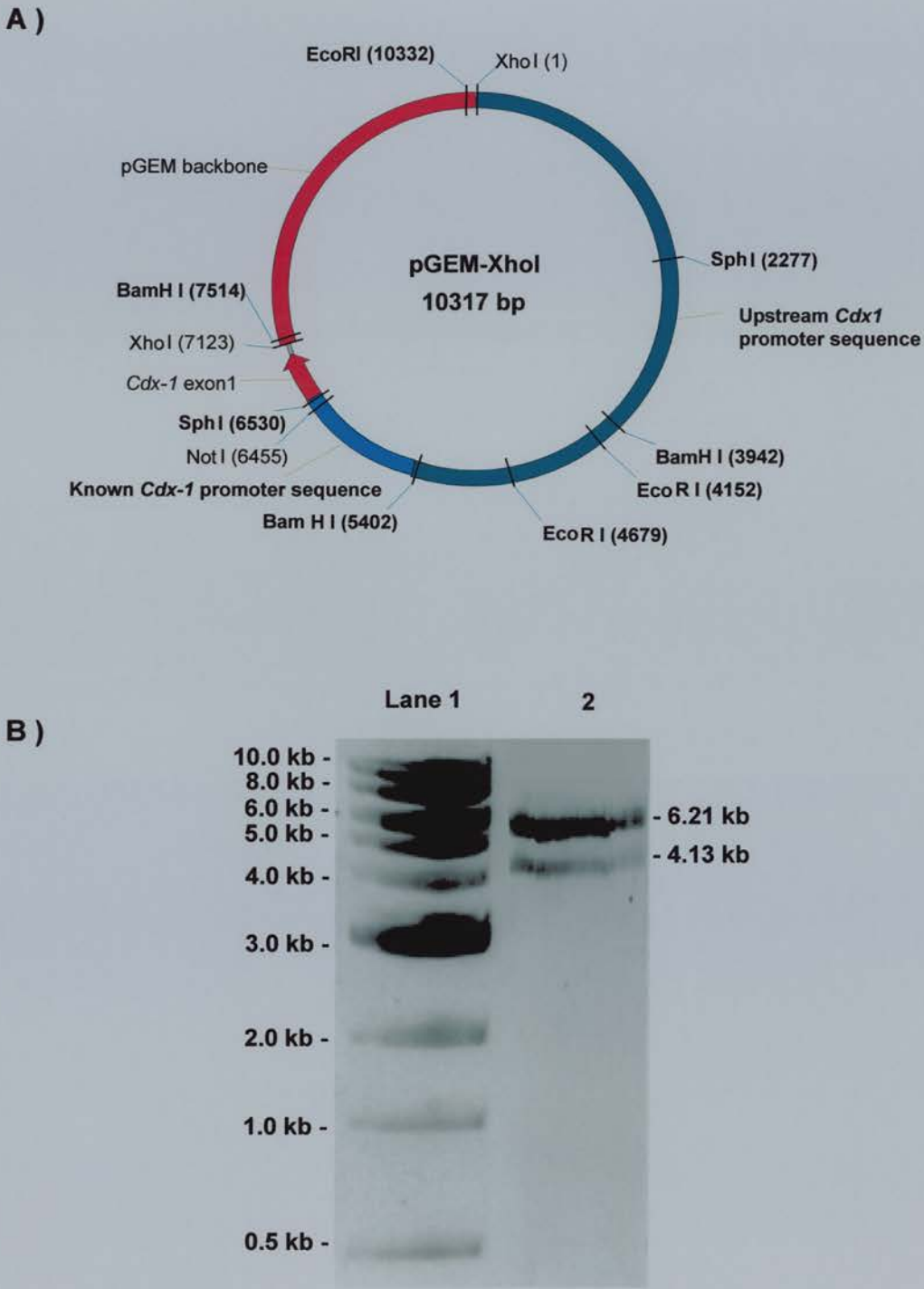


**Figure 4.4: pEGFP-1 reporter vector backbone**



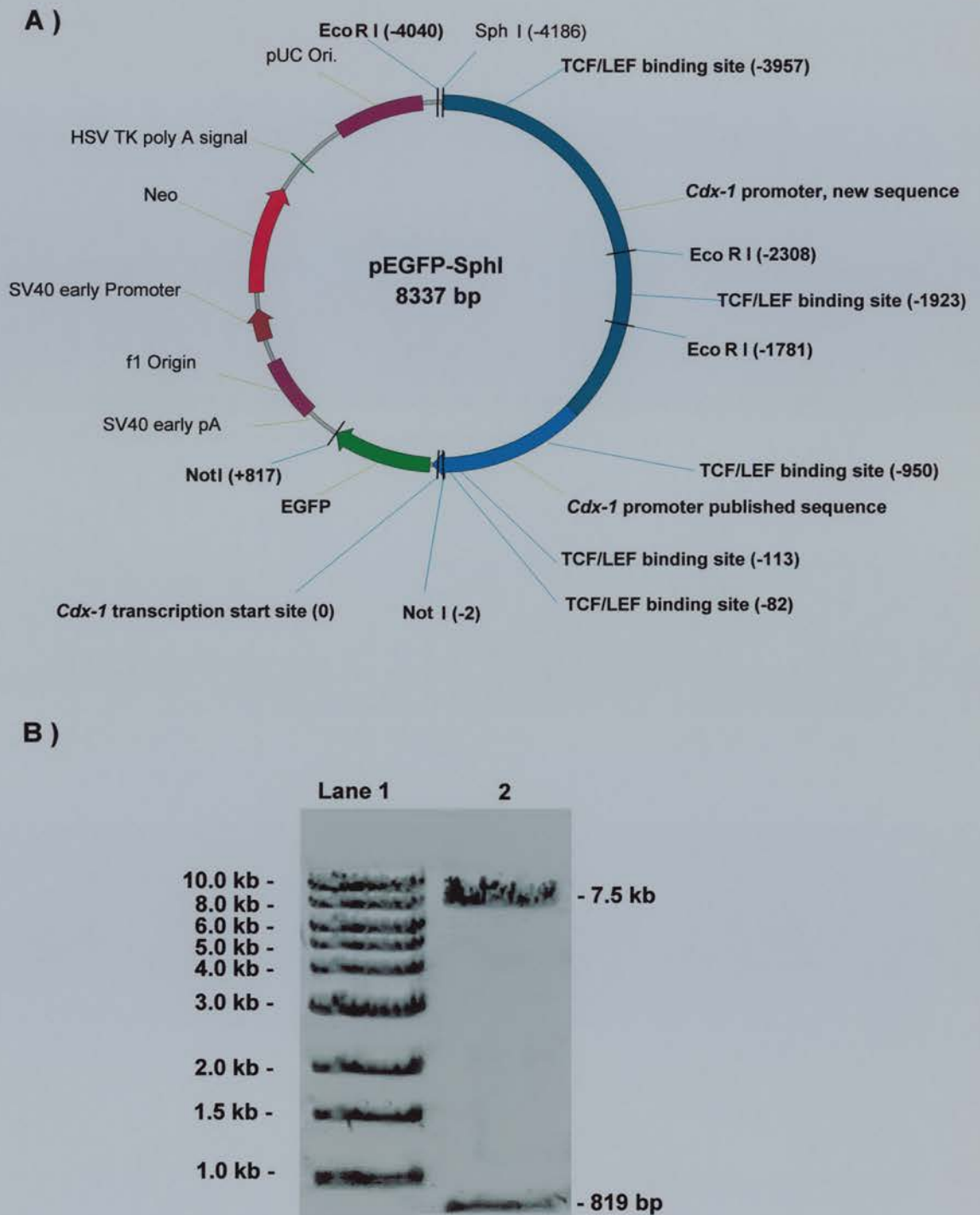
**Figure 4.4:** **A)** A vector schematic of the pEGFP-1 vector, encoding the Enhanced Green Fluorescent Protein (EGFP) sequence downstream of a multiple cloning site. There is no native promoter driving EGFP expression from this vector. The SmaI and EcoRI sites were used for the insertion of various portions of the murine *Cdx-1* promoter. The NotI site was used for analytic digests of the *Cdx-1*-EGFP constructs. **B)** Gel image of a SmaI digest of the pEGFP-1 plasmid. Lane 1: Kilobase ladder (NEB). Lane 2: SmaI digestion of pEGFP-1.

**Figure 4.5: Subcloning of pGEM-XhoI**



**Figure 4.5: A )** A vector schematic of the BAC 219 7.1 kb XhoI fragment cloned into the pGEM plasmid backbone. **B )** Digestion of the pGEM-XhoI plasmid with SphI liberates a 4.13 kb section of the subclone *Cdx-1* promoter sequence from the pGEM backbone. The digestion was supplemented with T4 DNA polymerase and 10mM nucleotides in order to generate blunt ended DNA molecules. Lane 1: Kilobase ladder (NEB). Lane 2: pGEM-XhoI SphI digestion showing 4.13kb *Cdx-1* promoter fragment.

**Figure 4.6: pEGFP-SphI reporter vector**



**Figure 4.6: A)** A vector schematic of the 4.13 kb SphI fragment of the pGEM-BAC 219 XhoI construct cloned into the SmaI site of pEGFP-1. The SphI fragment was blunted with T4 DNA polymerase prior to ligation into the blunt SmaI site. This section of the murine *Cdx-1* promoter contains 5 perfect TCF/LEF binding motifs. **B)** Gel electrophoresis of the *Cdx-1* pEGFP-Sph vector after restriction enzyme digestion. Lane 1: Kilobase ladder (NEB) Lane 2: NotI digest of the pEGFP-Sph vector showing the expected bands at 7.5 kb and 819 bp.

#### 4.4.2 Construction of pEGFP-LongEco and pEGFP-ShortEco

The 4.2kb SphI *Cdx-1* promoter fragment cloned into the pEGFP-1 vector contains 2 EcoRI sites, one at position -2308 and the other at position -1781 relative to the *Cdx-1* transcription start site. The EcoRI site at position -2308 lies downstream of the TCF/LEF site at position -3957 while the restriction site at position -1781 lies downstream of the TCF/LEF binding site at position -1923 relative to the *Cdx-1* transcription start site (see figure 4.1, 4.6a). The pEGFP-Sph vector contains an EcoRI site at 8191. The location of the EcoRI sites within pEGFP-Sph enables portions of the 4.2kb SphI *Cdx-1* promoter fragment to be removed; with a partial digestion and religation it was possible to generate the pEGFP-LongEco and pEGFP-ShortEco *Cdx-1* promoter plasmids simultaneously. The pEGFP-LongEco plasmid contains the EGFP coding sequence under the transcriptional control of 2381bp of the *Cdx-1* promoter, this 2.4kb contains 4 TCF/LEF binding motifs, the retinoic acid response element, the epithelial specific promoter element and the silencer/silencer blocking elements (see figure 4.7). pEGFP-ShortEco contains 1854bp of *Cdx-1* promoter containing the same elements although only 3 TCF/LEF binding motifs (see figure 4.8).

#### 4.4.3 Construction of pEGFP-SCdx

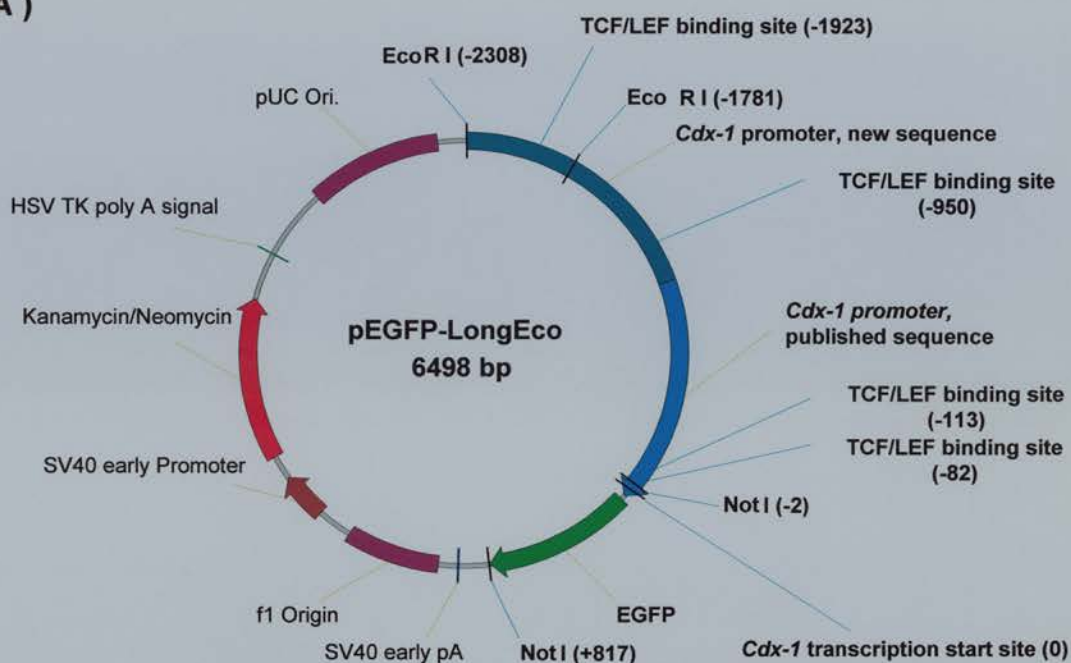
A lack of suitably compatible restriction site necessitated the use of PCR in order to generate a shorter *Cdx-1* promoter segment to drive EGFP transcription. The pBS-SCdx plasmid generated to provide homology arms for the BAC targeting vectors (see figure 3.16) contains a short PCR generated stretch of the *Cdx-1* promoter. Digestion of the pBS-SCdx plasmid with SphI supplemented with T4 DNA polymerase followed by digestion with EcoRI generates a 232bp fragment of the *Cdx-1* promoter containing the positive epithelial promoter element and 2 TCF/LEF binding sites. This fragment was cloned into the SmaI/EcoRI digested pEGFP-1 to generate the 4383bp pEGFP-SCdx (ShortCdx) plasmid (see figure 4.9). The *Cdx-1* promoter sequences contained within the four *Cdx-1*-EGFP reporter plasmids are summarised in figure 4.10.

The pCMV-GFP plasmid was used as a positive control for transfections and microinjections; the strong cytomegalovirus viral promoter efficiently drives transcription in a wide range of mammalian cells (see figure 4.11). GFP and EGFP are excited by similar wavelengths so detection of GFP in positive control experiments would affirm that the apparatus used to detect fluorescence were working properly.

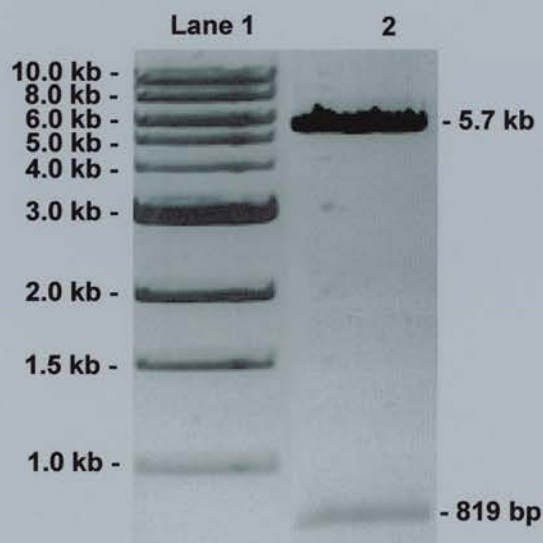


**Figure 4.7: pEGFP-LongEco reporter vector**

**A )**



**B )**

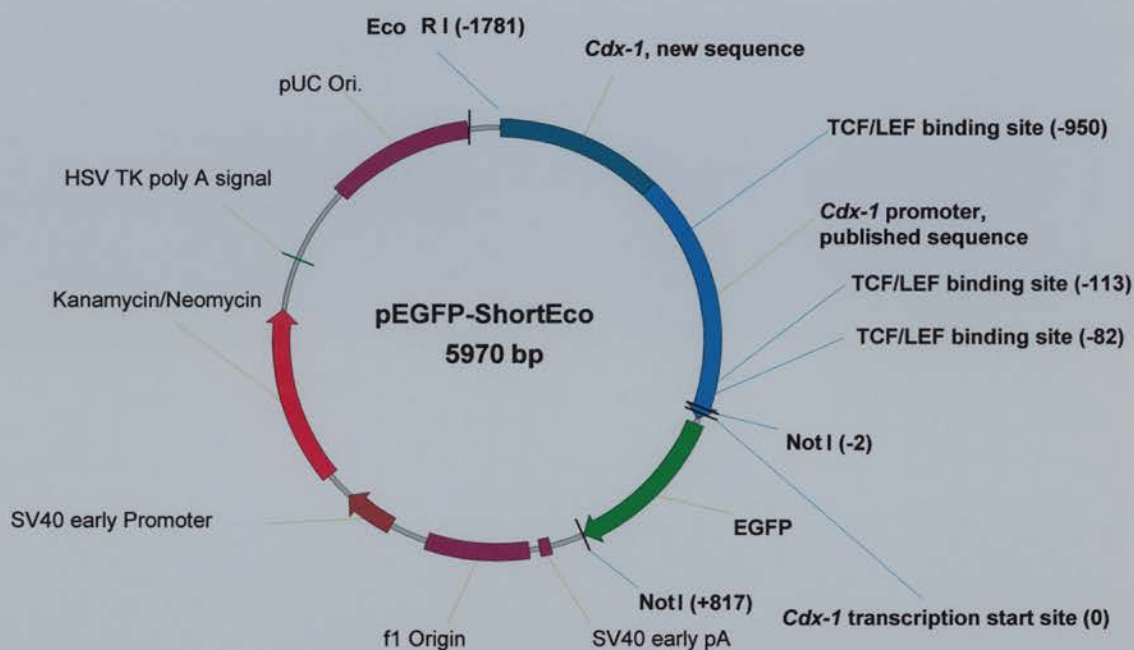


**Figure 4.7: A)** A vector schematic of the SphI fragment of the pGEM-BAC 219 XhoI construct cloned into the SmaI site of pEGFP-1 after the removal of a 1.8 kb EcoRI restriction fragment and subsequent religation. The resulting portion of the murine *Cdx-1* promoter contains 4 perfect TCF/LEF binding motifs. **B)** Gel electrophoresis of the pEGFP-LongEco vector after restriction enzyme digestion. Lane 1: Kilobase ladder (NEB) Lane 2: NotI digest of the pEGFP-LongEco vector showing the expected bands at 5.7 kb and 819 bp.

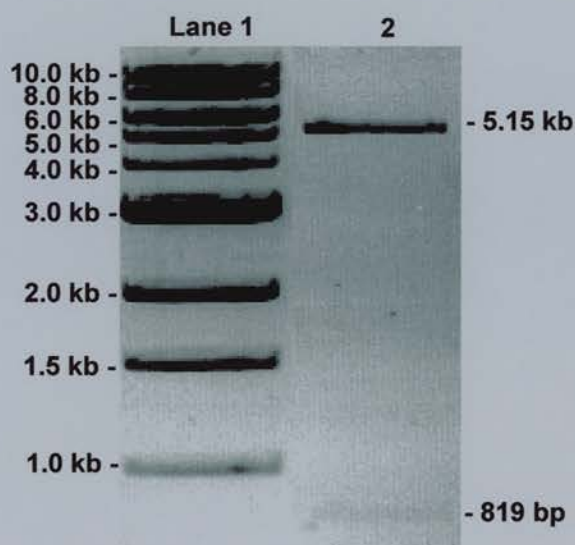


**Figure 4.8: pEGFP-ShortEco reporter vector**

**A )**

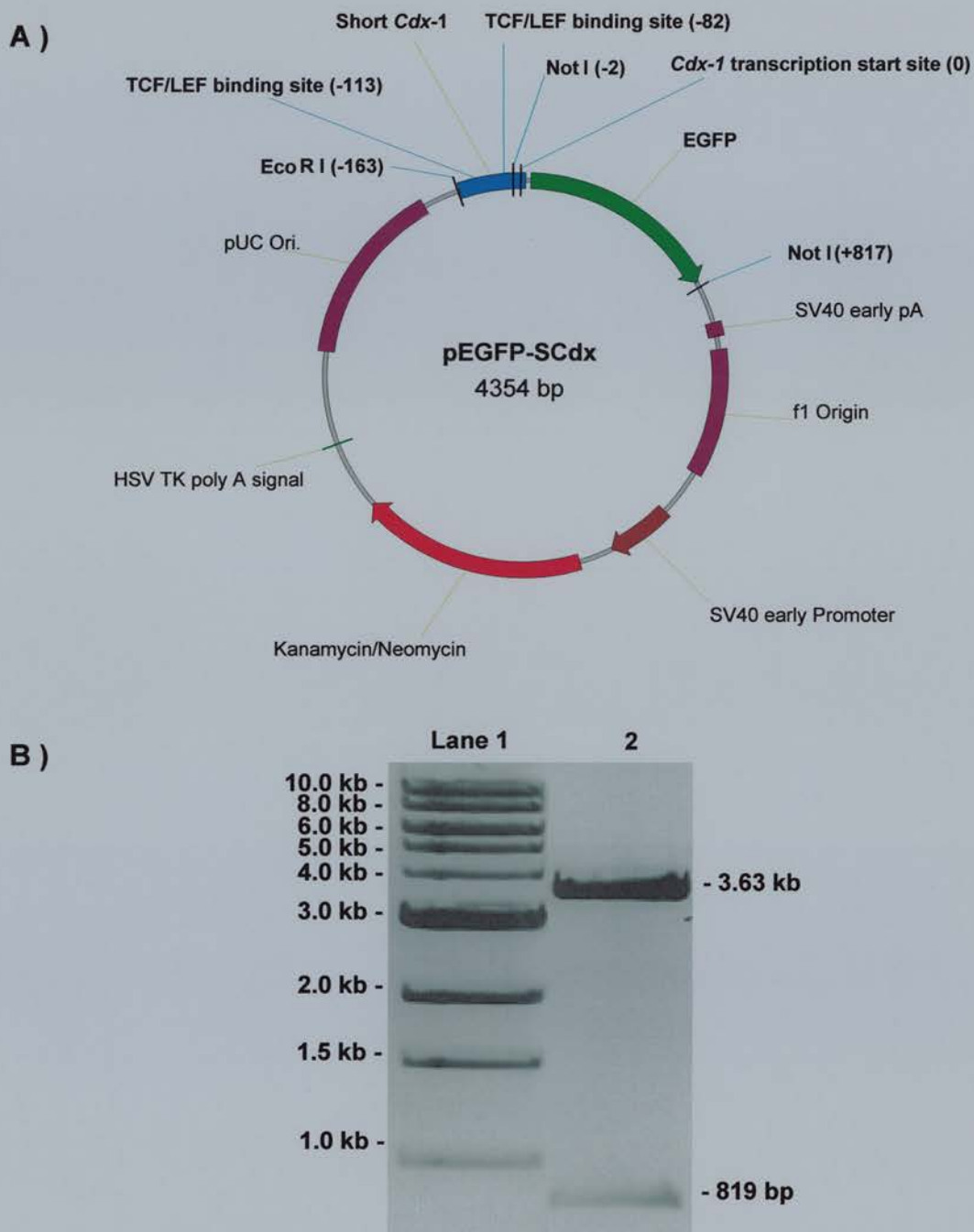


**B )**



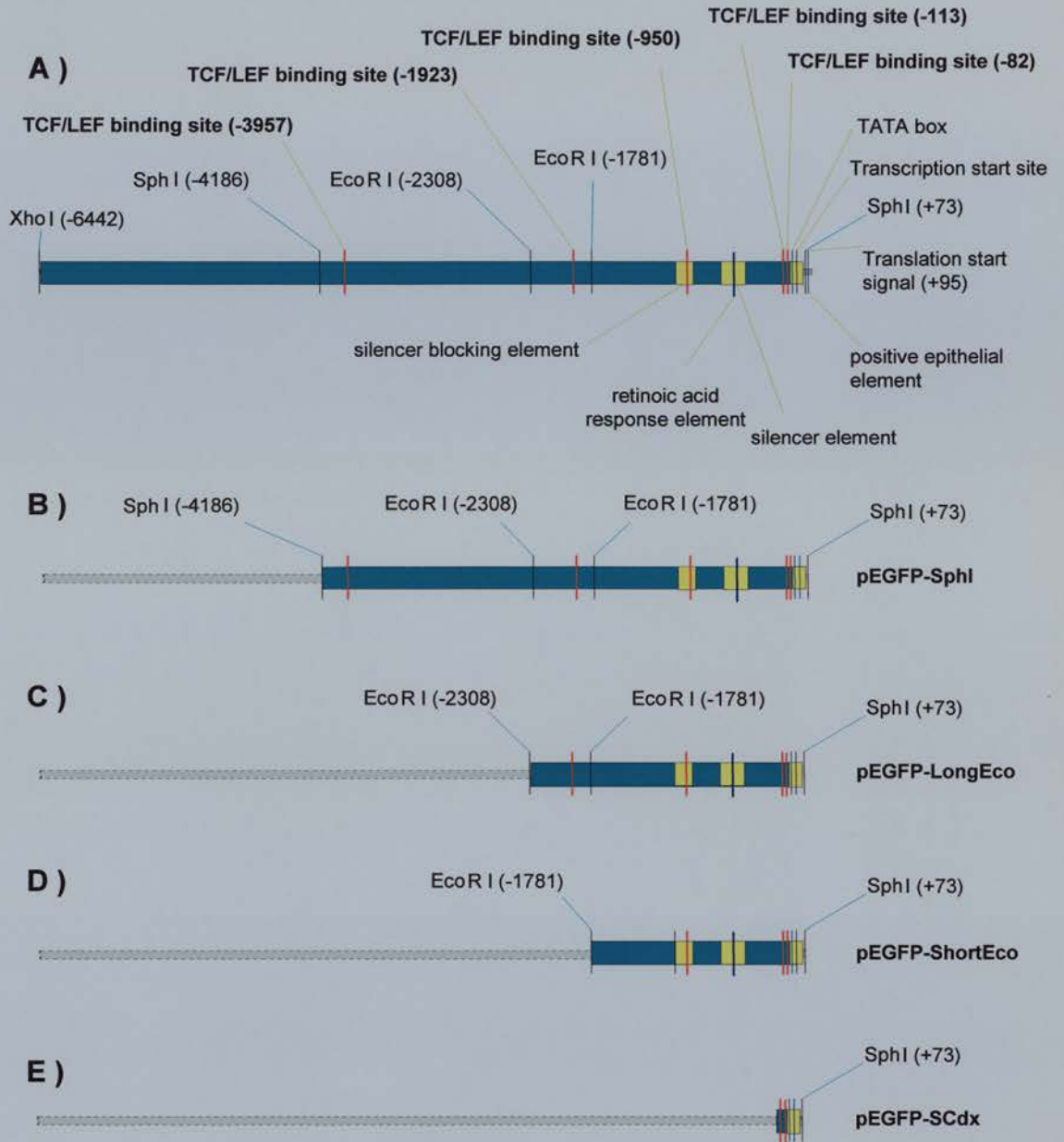
**Figure 4.8: A)** A vector schematic of the SphI fragment of the pGEM-BAC 219 XhoI construct cloned into the SmaI site of pEGFP-1 after the removal of a 2.3 kb EcoRI restriction fragment and subsequent religation. The resulting portion of the murine *Cdx-1* promoter contains 3 perfect TCF/LEF binding motifs. **B)** Gel electrophoresis of the pEGFP-ShortEco vector after restriction enzyme digestion. Lane 1: Kilobase ladder (New England Biolabs) Lane 2: NotI digest of the pEGFP-ShortEco vector showing the expected bands at 5.15 kb and 819 bp.

**Figure 4.9: pEGFP-SCdx reporter vector**



**Figure 4.9: A)** A vector schematic of the Short *Cdx-1* PCR cut with EcoRI/SphI and cloned into the EcoRI/SmaI sites of pEGFP-1. This portion of the murine *Cdx-1* promoter contains 2 perfect TCF/LEF binding motifs. **B)** Gel electrophoresis of the vector after restriction enzyme digestion. Lane 1: Kilobase ladder (New England Biolabs) Lane 2: NotI digest of the pEGFP-SCdx vector showing the expected bands at 3.63 kb and 819 bp.

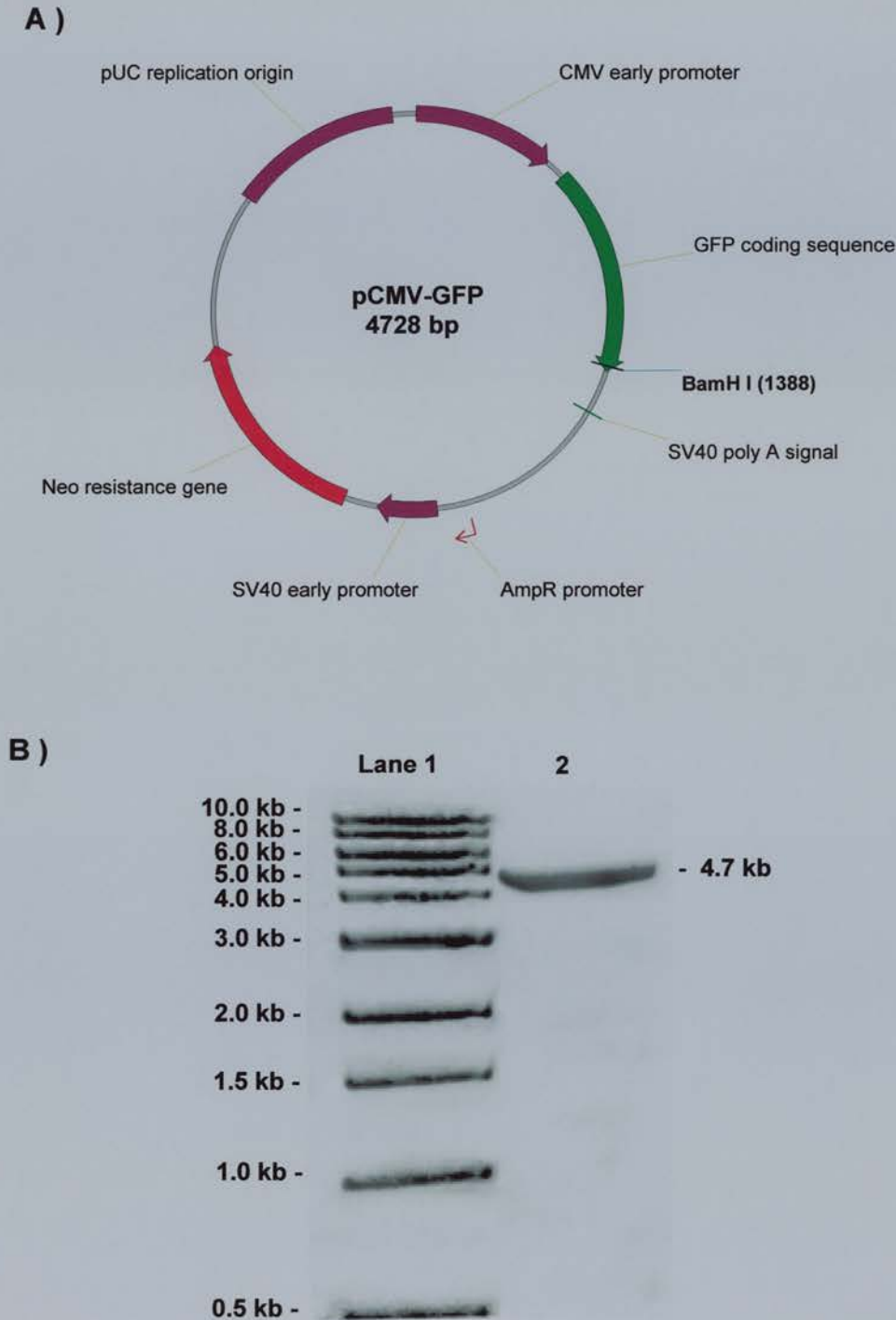
**Figure 4.10: Summary of *Cdx-1* reporter promoter sequences**



**Figure 4.10:** Schematics of the full *Cdx-1* XhoI promoter sequence showing known promoter elements and the shortened promoter fragments built into the pEGFP-1 promoterless reporter vector. **A)** 6.4 kb of sequence upstream of the start of *Cdx-1* transcription were isolated from BAC 219. The sequence contains a retinoic acid binding sequence, a silencer element, a silencer blocking element, a positive epithelial specific element and 5 perfect TCF/LEF binding motifs. **B)** The longest section of the isolated sequence cloned into the pEGFP-1 vector. The SphI digest fragment contains all the previously identified promoter elements as well as at least one newly identified TCF/LEF binding site. **C)** Truncated promoter sequence Long EcoRI. This sequence contains only 4 of the TCF/LEF binding sites while retaining the other identified elements. **D)** Truncated promoter sequence Short EcoRI. The sequence contains 3 TCF/LEF binding sites as well as all other identified promoter elements. **E)** Truncated promoter sequence Short *Cdx-1*. This sequence was generated by PCR from genomic template. It contains 2 TCF/LEF motifs and the epithelial specific element but lacks the retinoic acid response element and both the silencer and silencer blocking sequences.



**Figure 4.11: pCMV-GFP positive control reporter vector**



**Figure 4.11:** A) A vector schematic of the pCMV-GFP plasmid. This plasmid served as a positive control during the *Cdx-1* EGFP transient transfection experiments. B) Gel electrophoresis of the pCMV-GFP vector after restriction enzyme digestion. Lane 1: Kilobase ladder (New England Biolabs) Lane 2: BamHI digest of the pCMV-GFP plasmid showing the expected single band of 4.7 kb.

## 4.5 Introduction of pEGFP-*Cdx-1* plasmids into mammalian cells

To study the contribution of the *Cdx-1* promoter elements to overall levels of transcription the various pEGFP-*Cdx-1* vectors were introduced into a variety of mammalian cell lines. Initial experiments involved the human colorectal carcinoma derived cell lines HCT 116 and SW 480 and further experiments were performed using primary murine colonocytes and murine ES cells.

### 4.5.1 Microinjections

The use of microinjection as a means of introducing DNA into cells *in vitro* has some advantages over other methods. The technique can be very efficient, as robust cells will tolerate the procedure quite readily and up to 100% of cells within a particular region may be injected. Microinjection also allows for more consistent control over the amount of DNA entering any particular cell, as DNA concentrations are controlled for copy number and a precise amount of material is injected into each cell. The relative consistency in the number of DNA molecules introduced into a cell as well as the high efficiency make comparison of levels of transfection from the reporter constructs more immediately quantitative, at least on a relative scale. One disadvantage of microinjection is the relative harshness of the procedure, fragile cell lines or primary cultures may not be able to withstand even the most mild injection pressures and may rupture immediately or die soon after microinjection. Additionally cell morphology can make microinjection very difficult; rounded cells tend to be difficult to inject while flat epithelial cells are significantly easier to successfully microinject. As SW 480 cells tend towards a more rounded morphology microinjections were carried out on HCT 116 cells and primary murine colonocytes as their morphology makes them considerably easier to accurately inject.

HCT 116 and primary murine colonocytes were plated into 5cm culture dishes in preparation for microinjections as the angle of the injection needle and apparatus makes injecting into wells awkward. The human cell line was allowed to reach 75% confluency prior to injecting while the murine colonocytes were injected 3-5 days after isolation and plating. Control and reporter plasmid preparations were

diluted to achieve consistent plasmid copy number per unit volume in injection buffer (see appendix 1). In some applications a fluorescent tracer dye can be added to the injection buffer in order to check that the injection has succeeded in introducing material into the cell: in this case, as the dye can persist for 24-48 hours depending on the cell injected and may mask the fluorescence (or lack thereof) generated by transcription from the reporter plasmids, the injections were carried out without adding the tracer dye. The relative plasmid size and final DNA concentration in the injection medium for both reporter and control constructs are given in table 4.1

**Table 4.1: Plasmid DNA concentrations for microinjections**

Plasmid	Size (bp)	DNA concentration in injection medium
pCMV-GFP (positive control)	4728	125ng/μl
pEGFP-1 (negative control)	4151	109ng/μl
pEGFP-SCdx	4354	115ng/μl
pEGFP-ShortEco	5970	158ng/μl
pEGFP-LongEco	6498	172ng/μl
pEGFP-Sph	8211	217ng/μl

Microinjections were preformed using the apparatus and protocols previously described (see section 2.7.1) and the injected cells monitored at intervals under normal phase contrast and UV illumination for up to 48 hours following the introduction of DNA.

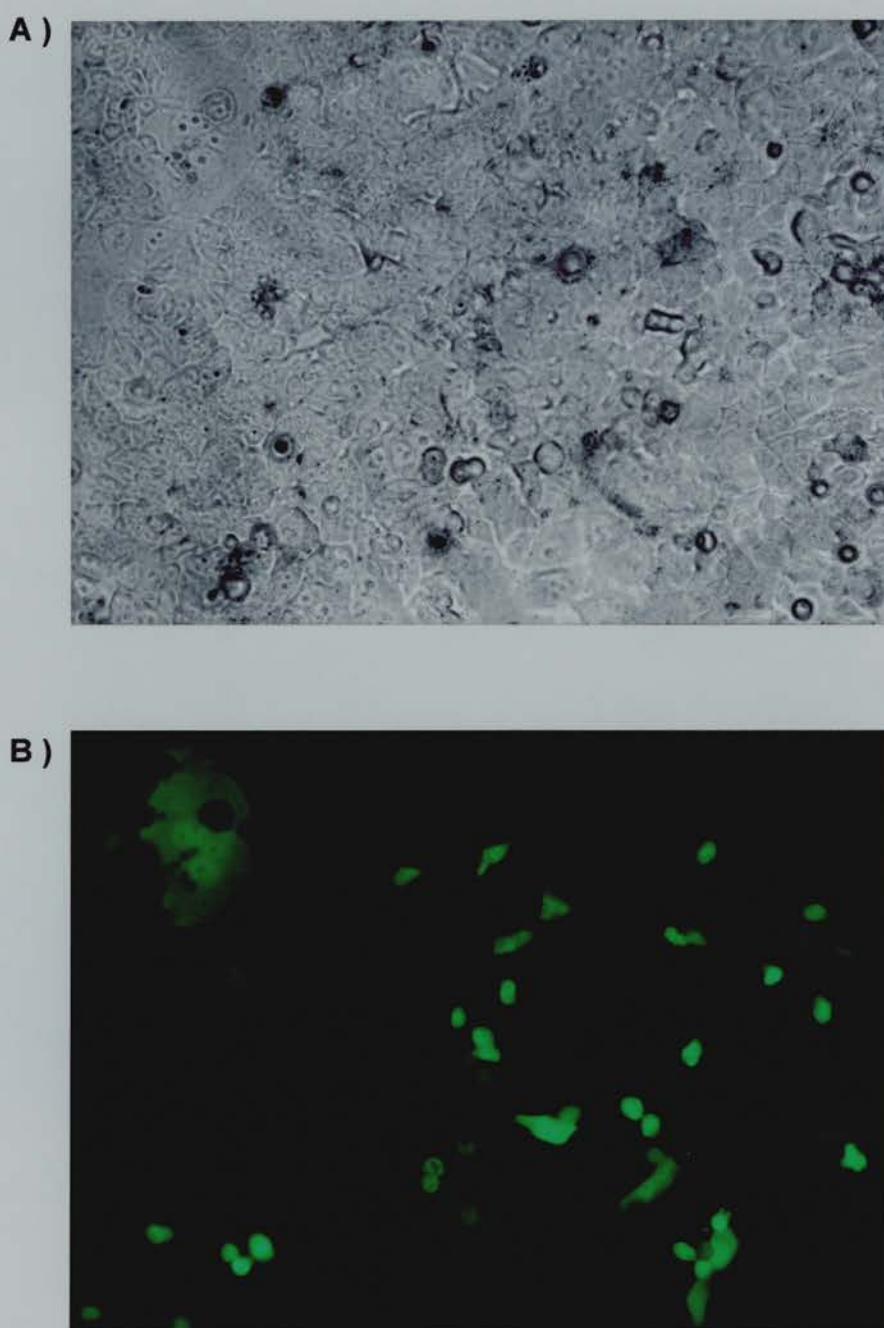
While the HCT 116 human cell line proved able to tolerate the microinjection procedure, the primary murine colonocytes seemed too fragile to undergo injection and survive for the periods needed to observe possible low level EGFP expression from the reporter plasmids. HCT 116 cells injected with the positive control pCMV-GFP reporter plasmid were observed to fluoresce when UV illuminated at intervals over the full monitoring period, indicating that the injected cells were surviving the procedure and persisting in culture (see figure 4.12). The HCT 116 cells microinjected with *Cdx-1*-EGFP reporter plasmids failed to show any signs of fluorescence over the period of observation (48



hours). That the cells appeared to survive the microinjection procedure and remain healthy indicated that the failure to observe fluorescence was not due to cell death but instead due to insufficient expression from the reporter plasmids.

Primary murine colonocytes were observed to fluoresce following injection with the pCMV-GFP positive control plasmid beginning 3 hours post injection but these cells appeared to die by 6-8 hours after introduction of plasmid DNA. This short time period appears to be insufficient to observe EGFP at the levels of transcription supported by the *Cdx-1* fragments built into the reporter construct.

**Figure 4.12: Microinjection of HCT 116 cells**



**Figure 4.12: A)** Phase contrast image of HCT 116 human colorectal cell line (200x). **B)** UV illuminated image of HCT 116 cells microinjected with the positive control pCMV-GFP plasmid 48 hours post injection (200x). HCT 116 cells injected with Cdx1-EGFP reporter plasmids showed no fluorescence above background in the 48 hours following microinjection.

#### 4.5.2 Transfections

In addition to the microinjection procedures, a variety of cell types were liposome transfected with the control and reporter constructs. While transfection efficiencies can vary between cell types and from one plasmid to another making accurate quantitative comparisons more difficult, the procedure is relatively mild and fragile cells such as primary murine colonocytes or ES cells may survive long enough to show evidence of *Cdx-1* promoter driven EGFP transcription. The ease with which the transfection reactions can be carried out as well as the relatively small volumes and numbers of cells involved made it possible to run a parallel transfection using the same plasmids but with the cell medium being supplemented with 1 $\mu$ M retinoic acid (RA). As mentioned previously, various concentrations of retinoic acid have been reported to upregulate expression from the *Cdx-1* promoter in murine embryos and embryonalcarcinoma cell lines. If the levels of transcription driven by the *Cdx-1* EGFP reporter constructs containing the retinoic response element (RARE) incorporated into the EGFP reporter plasmids were very low, the addition of RA to the medium may bring the levels of fluorescence into the detectable range.

HCT 116, SW 480, murine ES cells and murine colonocytes were seeded into 24 well plates in preparation for transfection experiments. As in the microinjection procedure the cell lines were allowed to reach 75% confluency prior to introduction of plasmid DNA while the colonocytes were transfected 3-5 days post plating. Transfections were carried out using FuGENE reagent (Roche, see chapter 2 for protocol) and the optimal ratio of reagent to DNA determined to be 6 $\mu$ l FuGENE:1 $\mu$ g DNA by transfection with the positive control plasmid pCMV-GFP. This transfection ratio appeared to be optimal across all three cell types although the amount of reagent:DNA added to a well varied from 12 $\mu$ l:2 $\mu$ g for the colonocytes to 6 $\mu$ l:1 $\mu$ g for the cell lines, the increased amount of transfection mixture added to the colonocyte preparations did not appear to have any toxic effect on the cells. The transfection efficiencies for the human derived cell line HCT 116 and the ES cells were approximately 30%, while <10% of primary murine colonocytes and SW 480 cells expressed GFP after transfection.

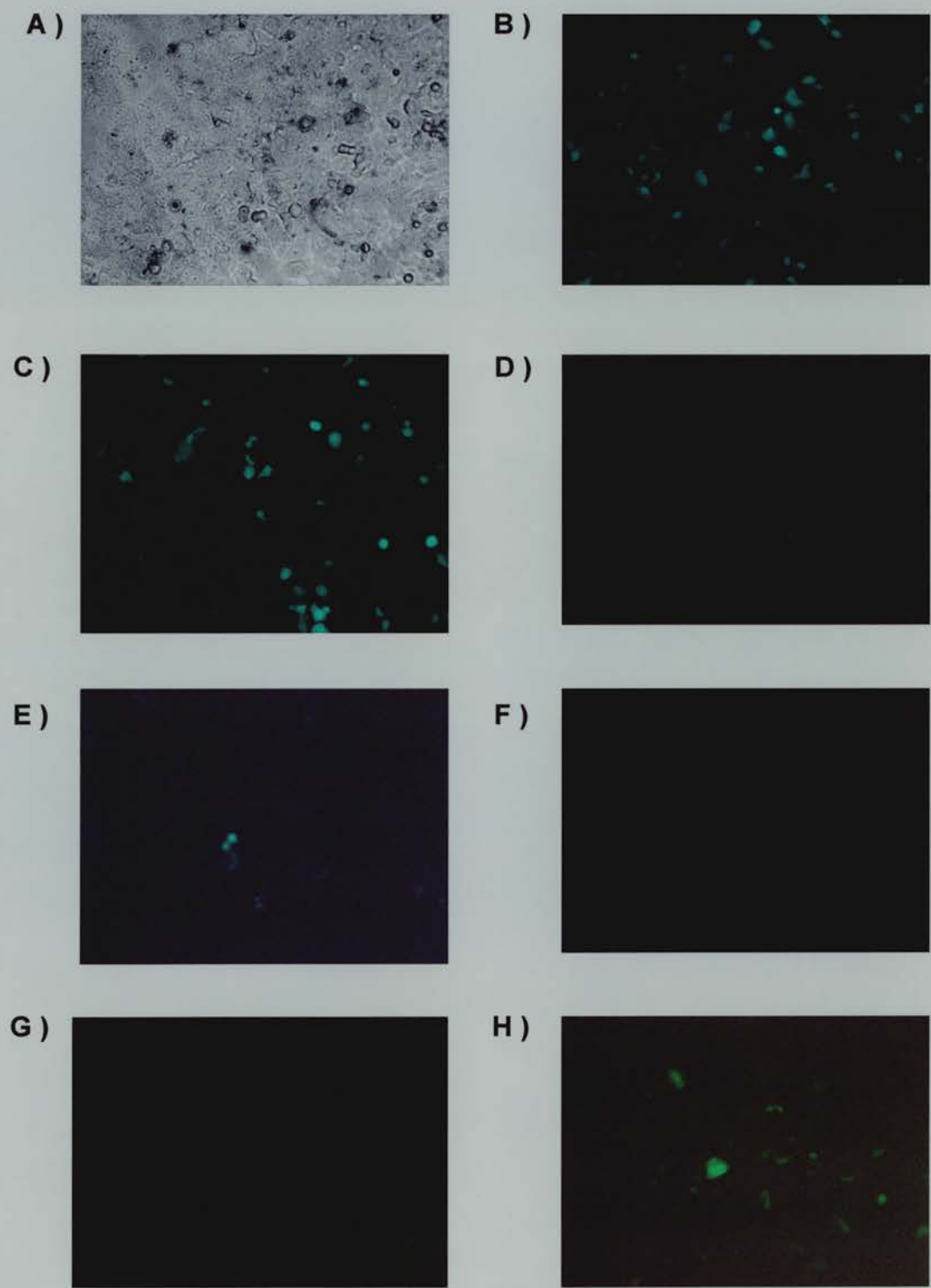
Following administration of the transfection mixture the cells were monitored at intervals for 48 hours under normal phase contrast and UV illumination.

#### 4.5.2.1 HCT 116 cell line

HCT 116 cells began to show evidence of fluorescence in pCMV-GFP transfected cells within 3 hours of application of DNA and continued in a constant manner for at least 48 hours (see figure 4.13b and 4.13c) while *Cdx-1*-reporter transfected populations began to show signs of fluorescence approximately 36 hours following the introduction of DNA and continuing for at least a further 12 hours (see figure 4.13). While relatively widespread fluorescence was observed from the positive control pCMV-GFP construct the fluorescence observed from *Cdx-1*-promoter driven constructs was considerably more dispersed. Small patches of cells were observed showing fluorescence in those wells transfected with the minimal *Cdx-1*-reporter, pEGFP-SCdx (see figure 4.13e). As mentioned previously this plasmid contains only 2 TCF/LEF binding motifs and the positive epithelial control element. Transfections with the pEGFP-ShortEco and pEGFP-LongEco reporter plasmids failed to produce fluorescence in the transfected HCT 116 cells in the 48 hours following transfection and, again, the addition of 1 $\mu$ M retinoic acid to the culture medium failed to produce any visible increase in fluorescence, despite the fact that these plasmids contain the RARE. HCT 116 cells transfected with the pEGFP-Sph reporter construct showed relatively widespread and intense fluorescence beginning 36 hours post-transfection (see figure 4.13h). While this plasmid also contains the retinoic acid response element, no change in fluorescence was observed in cells incubated with 1 $\mu$ M retinoic acid.

**Figure 4.13:** **A)** Phase contrast image of HCT 116 human colorectal cell line (200x). **B)** UV illuminated image of HCT 116 cells transfected with the positive control pCMV-GFP plasmid 8 hours post transfection (200x). **C)** HCT 116 cells 48 hours after transfection with pCMV-GFP (200x). **D)** HCT 116 cells 48 hours after transfection with the promoterless pEGFP-1 plasmid (200x). **E)** HCT 116 cells 48 hours after transfection with the pEGFP-SCdx plasmid showing low levels of fluorescence (200x). **F)** Cells transfected with the pEGFP-ShortEco reporter. **G)** HCT 116 cells transfected with the pEGFP-LongEco plasmid (200x). **H)** HCT 116 cells 48 hours after transfection with the pEGFP-SphI plasmid showing relatively high levels of fluorescence (200x). The cells transfected with pEGFP-ShortEco and pEGFP-LongEco showed no fluorescence over background levels.

Figure 4.13: Transfections of the human derived HCT 116 colorectal cell line



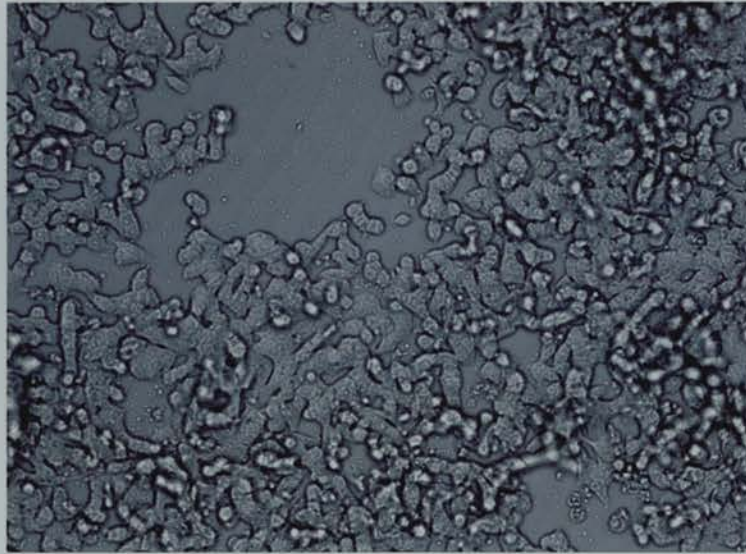


#### 4.5.2.2 SW 480 cell line

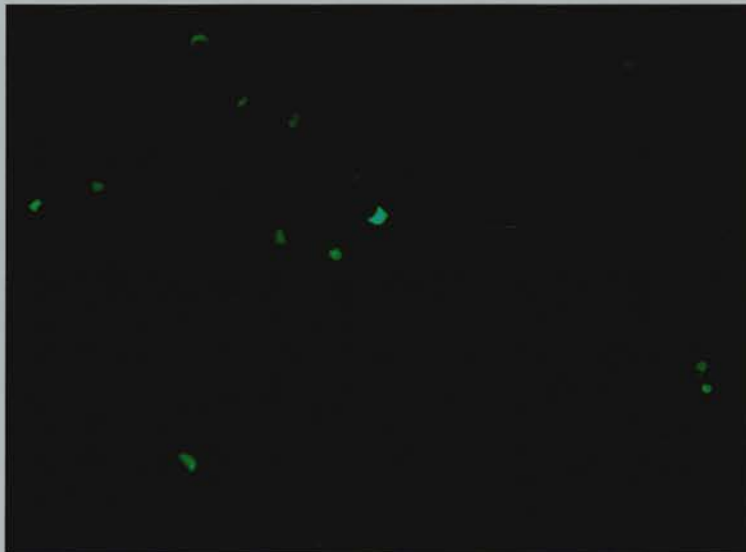
SW 480 cells, when compared to HCT 116 cells, showed a reduced transfection efficiency and were very slow growing. Transfection with pCMV-GFP resulted in <10% of cells showing significant fluorescence 36 hours after transfections (see figure 4.14b). SW 480 cells also differ from HCT 116 cells in morphology; they tended to remain relatively rounded rather than showing the more regular epithelial morphology of the HCT 116s (see figures 4.13a and 4.14a for comparison). This rounded morphology made detecting low levels of fluorescence difficult as rounded cells appear to “lens” and auto-fluoresce significantly more than flat cell types, perhaps overwhelming GFP/EGFP emission. Transfection with the *Cdx-1*-reporter constructs did not appear to result in fluorescence significantly greater than the background observed in the pEGFP-1 promoterless negative control transfections, again regardless of RA application.

**Figure 4.14: Transfection of SW 480 cells**

**A )**



**B )**



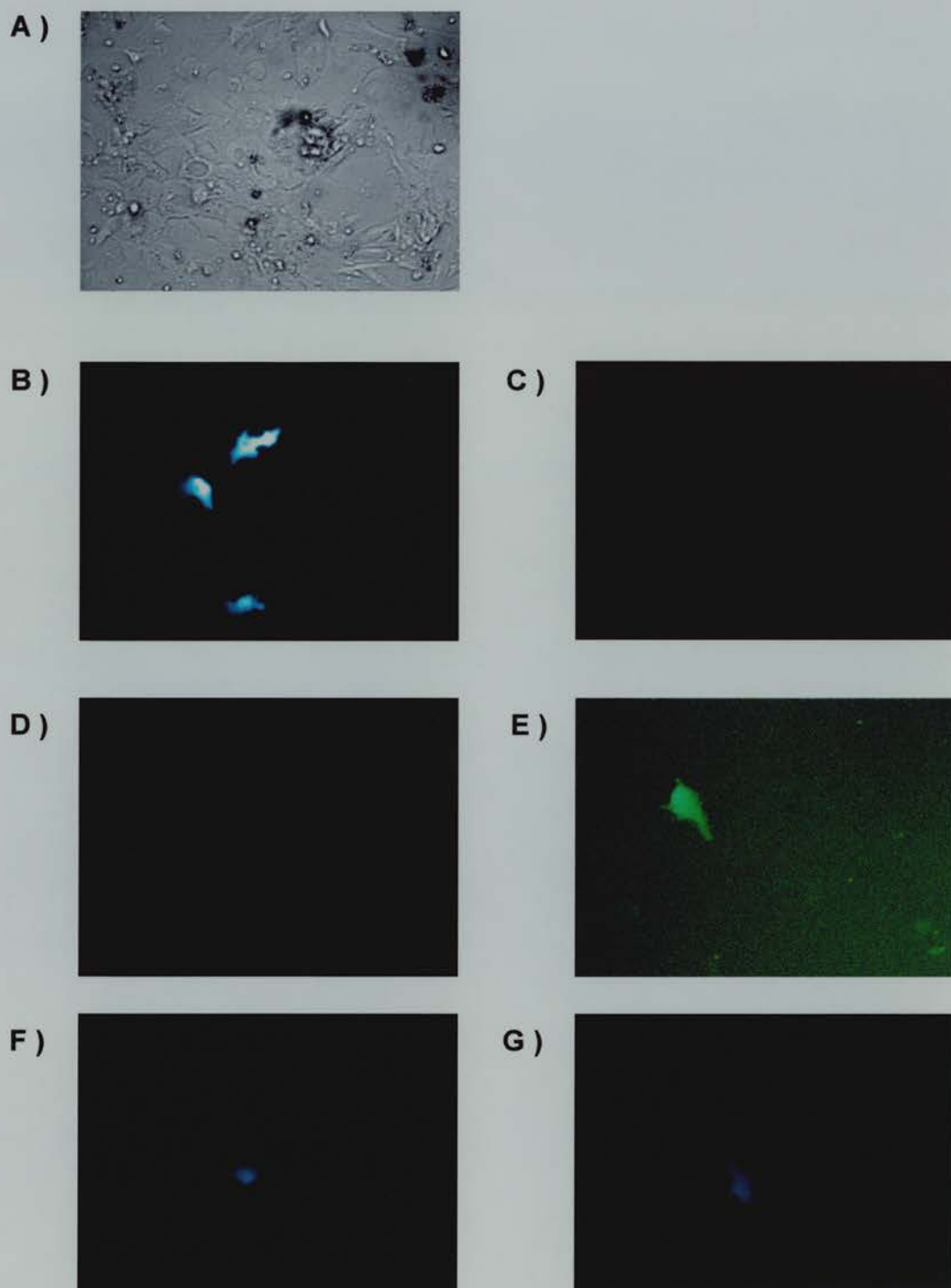
**Figure 4.14: A)** Phase contrast image of SW 480 human colorectal cell line (200x). **B)** UV illuminated image of SW 480 cells transfected with the positive control pCMV-GFP plasmid 48 hours post transfection (200x). No EGFP fluorescence was observed in SW 480 cells transfected with the *Cdx-1*-EGFP constructs

#### **4.5.2.3 Primary murine colonocytes**

Transfection of colonocyte preparations achieved a variable efficiency, with maximum transfection levels being ~10%. This variability may have been in part due to the variability in plating density from one preparation to another and also may have been influenced by the presence of fibroblasts within the platings. A relatively large amount of debris was also present within the colonocyte preparations and this debris showed a tendency to autofluoresce. Despite the variability of transfections and the presence of autofluorescing debris, it was possible to identify colonocytes expressing the pCMV-GFP reporter plasmid and to distinguish them from the fibroblast contamination based on morphology. Colonocytes containing the pCMV-GFP plasmid showed relatively high levels of fluorescence; these fluorescing cells tended to occur in groups within islands formed when the isolated colonic crypts adhered to the tissue culture vessel (see figure 4.15b). Transfection with the promoterless pEGFP-1 plasmid failed to produce fluorescence above the relatively high background within the 48 hours of observation, as did transfection with the pEGFP-SCdx reporter construct (see figure 4.15c and 4.15d). Very low levels of fluorescence were observed in the colonocyte preps transfected with the pEGFP-ShortEco reporter while higher, but still faint, fluorescence was achieved in isolated cells following transfection with either the pEGFP-LongEco or pEGFP-Sph reporter constructs (see figure 4.15e and 4.15f).

**Figure 4.15:** **A)** Phase contrast image of murine primary colonocytes (200x). **B)** UV illuminated image of murine colonocyte cells transfected with the positive control pCMV-GFP plasmid 48 hours post transfection (200x). **C)** Murine colonocytes 48 hours after transfection with the promoterless pEGFP-1 vector. The cells transfected with the negative control pEGFP-1 plasmid showed no fluorescence over background levels (200x). **D)** Murine colonocytes 48 hours after transfection with pEGFP-SCdx, no fluorescence above background was detectable (200x). **E)** Murine colonocytes 48 hours after transfection with the pEGFP-ShortEco plasmid; the image has been enhanced to show the very low level of fluorescence presence in a small number of cells (200x). Fluorescence was lower than that observed in transfections with pEGFP-LongEco or pEGFP-SphI. **F)** Murine colonocytes cells 48 hours after transfection with the pEGFP-LongEco plasmid showing low levels of fluorescence (200x). **G)** Murine colonocytes cells 48 hours after transfection with the pEGFP-SphI plasmid showing slightly higher levels of fluorescence (200x).

**Figure 4.15: Transfections of murine primary colonocytes**



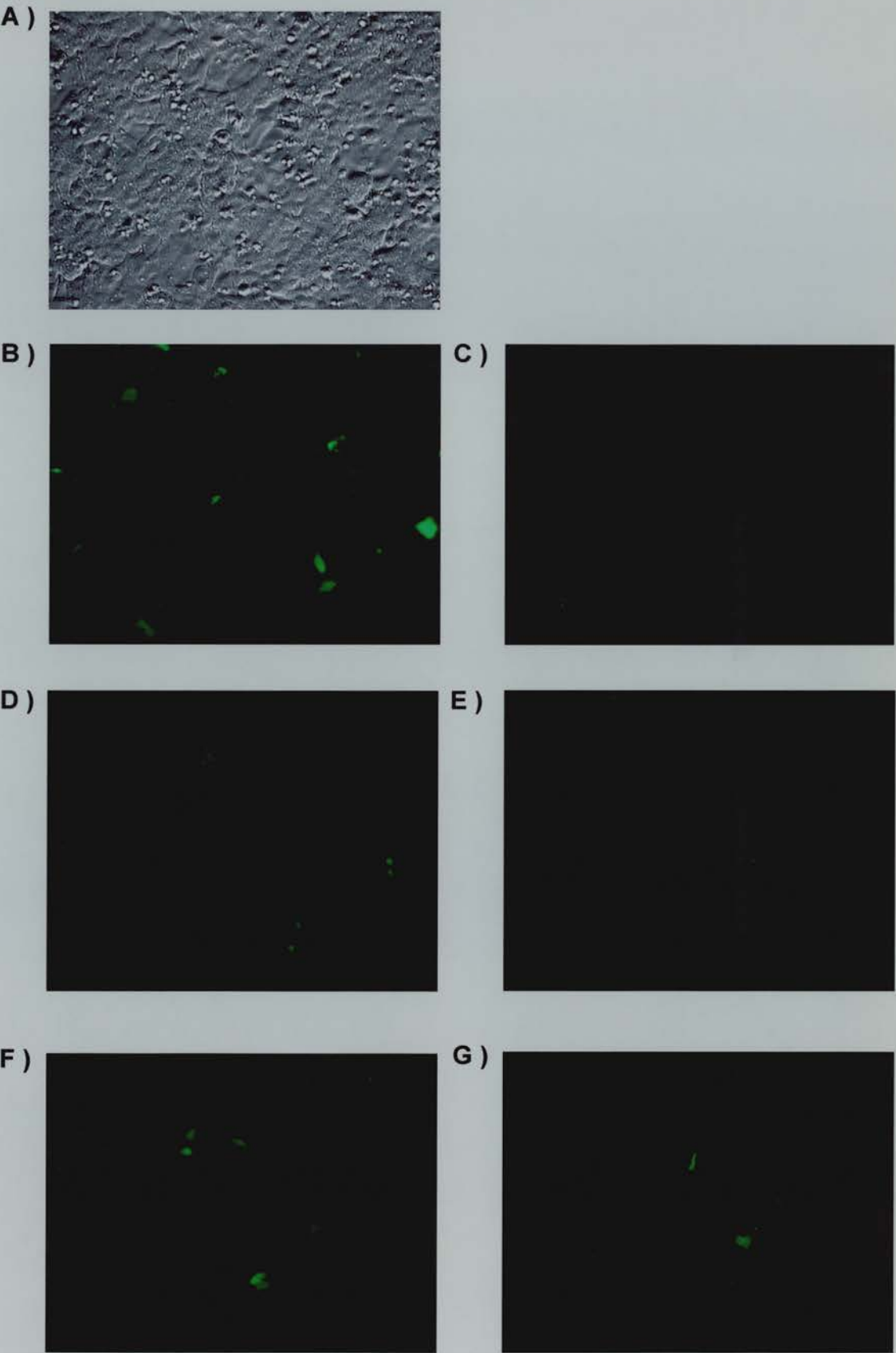
#### 4.5.2.4 Murine HM-1 ES cells

Liposome mediated transfection of murine ES cells with the pCMV-GFP plasmid resulted in an efficiency of ~30%, comparable to that achieved in the HCT 116 cell line. The appearance of fluorescence from the positive control transfection occurred within 3 hours of transfection, again comparable with the time frame seen in other cell types. The appearance of fluorescence following transfection of the *Cdx-1*-EGFP reporter plasmids, however, occurred much more quickly than in the other cell type, with visible levels of fluorescence appearing as little as 12 hours following DNA application (see figure 4.16). Additionally, the intensity of fluorescence provoked by the reporter plasmids was much higher in the ES cells. The level of fluorescence increased from 12 hours to 24 hours and remained relatively constant for the period up to 48 hours (see figure 4.16 and 4.17). pEGFP-SCdx generated a relatively strong fluorescence in ES cells, while the pEGFP-ShortEco reporter failed to produce any obvious fluorescence above background levels (see figure 4.17c and 4.17d). The pEGFP-LongEco produced the most intense fluorescence, while the pEGFP-Sph reporter provoked an intermediate fluorescence response (see figure 4.17e and 4.17f). The addition of 1 $\mu$ M RA to the culture medium for the 12 hours following DNA transfection did not appear to have any upregulating effect of transcription from the *Cdx-1*-EGFP reporter plasmids, even with those reporter constructs which contained the reported RARE. The levels of fluorescence actually appeared to be reduced when compared to HM-1 cells which had not been treated with retinoic acid (see figure 4.18).



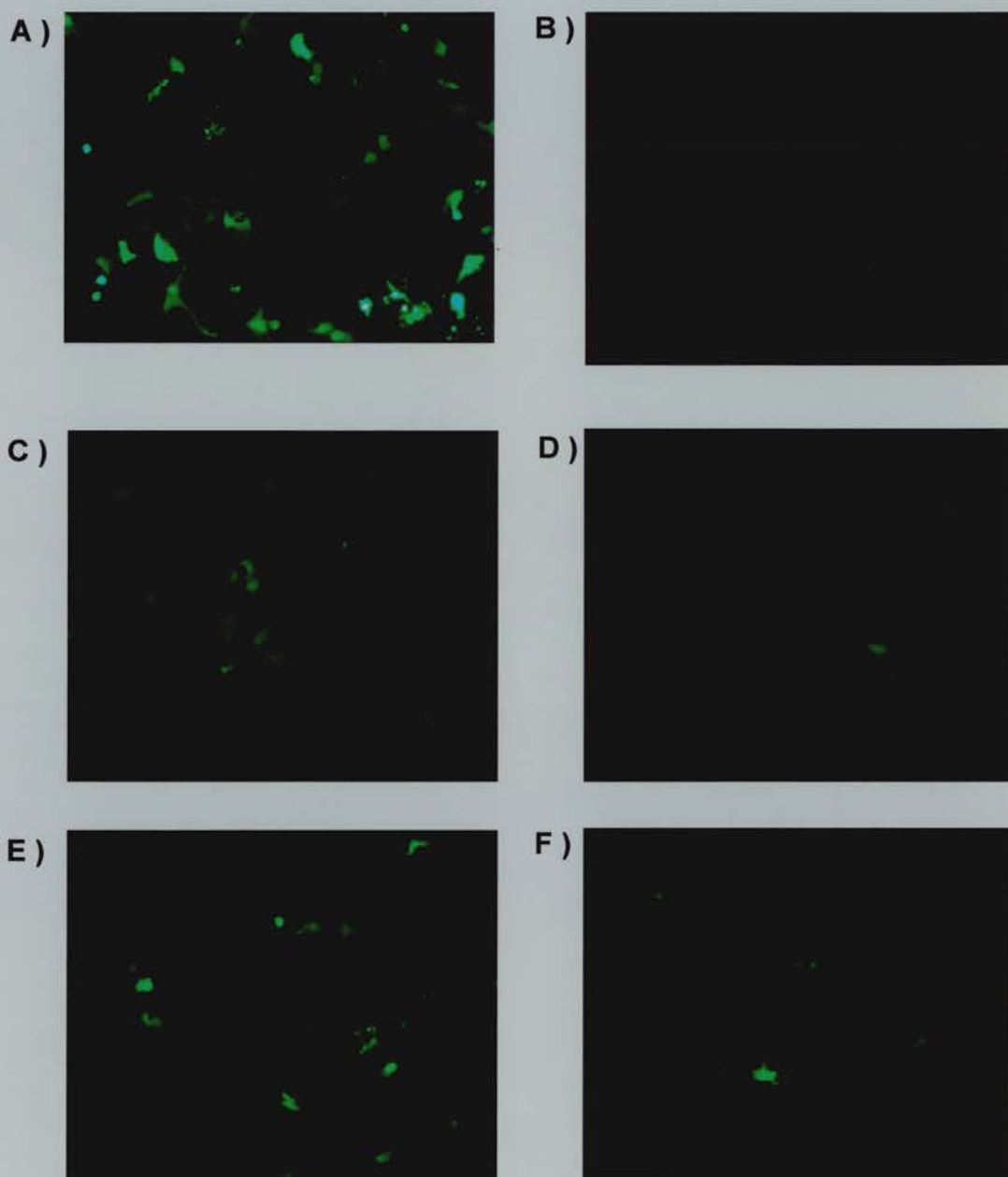
**Figure 4.16:** **A)** Phase contrast image of the HM-1 murine stem cell line (200x). **B)** UV illuminated image of HM-1 cells transfected with the positive control pCMV-GFP plasmid 12 hours post transfection (200x). **C)** HM-1 cells 12 hours after transfection with the promoterless pEGFP-1 plasmid (200x). **D)** HM-1 cells 12 hours after transfection with the pEGFP-SCdx plasmid showing low levels of fluorescence (200x). **E)** HM-1 cells 12 hours after transfection with the pEGFP-ShortEco plasmid showing no fluorescence above background (200x). **F)** The HM-1 cells transfected with pEGFP-LongEco showed relatively high levels of fluorescence (200x). **G)** HM-1 cells transfected with pEGFP-Sph again showing relatively high levels of fluorescence (200x).

**Figure 4.16: Activity of Cdx1-EGFP reporter constructs in ES cells  
12 hours post-transfection**



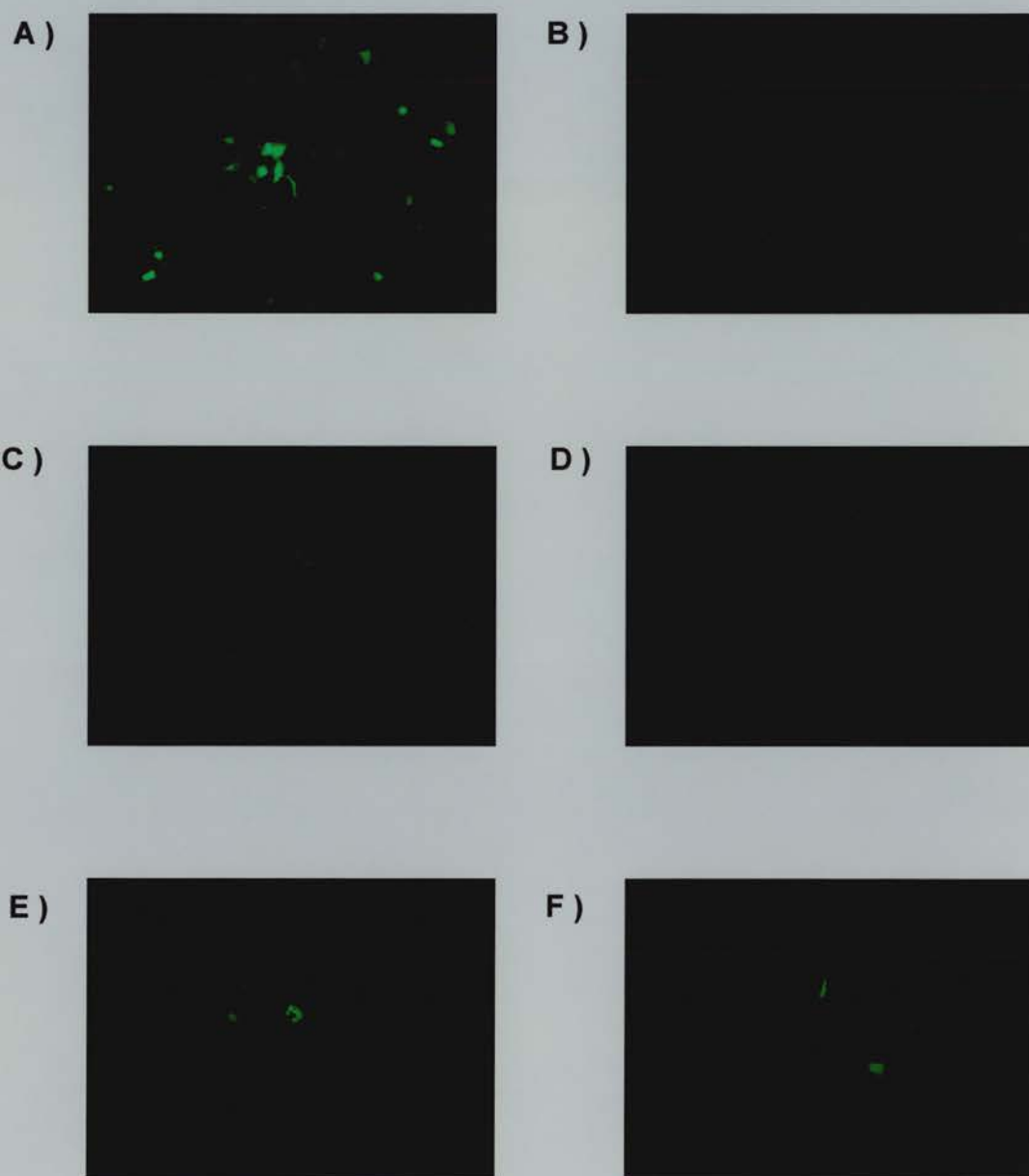
**Figure 4.17:** HM-1 cells 24 hours post-transfection. **A)** UV illuminated image of HM-1 cells transfected with the positive control pCMV-GFP plasmid 24 hours post transfection (200x). The fluorescence is comparable and perhaps stronger than that observed at 12 hours. **B)** HM-1 cells 24 hours after transfection with the promoterless pEGFP-1 plasmid again showing very little background fluorescence (200x). **C)** HM-1 cells 24 hours after transfection with the pEGFP-SCdx plasmid showing higher levels of fluorescence than seen at 12 hours (200x). **D)** HM-1 cells 24 hours after transfection with the pEGFP-ShortEco plasmid again showing no fluorescence above background (200x). **E)** The HM-1 cells transfected with pEGFP-LongEco showing levels of fluorescence relatively higher than those seen at 12 hours (200x). **F)** HM-1 cells transfected with pEGFP-Sph again showing moderate levels of fluorescence (200x).

**Figure 4.17: Activity of Cdx1-EGFP reporter constructs in ES cells  
24 hours post-transfection**



**Figure 4.18:** Effect of incubation with 1 $\mu$ M retinoic acid on expression from the Cdx1-EGFP reporter plasmids in HM-1 ES cells. **A)** UV illuminated image of HM-1 cells transfected with the positive control pCMV-GFP plasmid 12 hours post transfection (200x). The fluorescence is comparable to that observed without retinoic acid. **B)** HM-1 cells 12 hours after transfection with the promoterless pEGFP-1 plasmid again showing very little background fluorescence (200x). **C)** HM-1 cells 12 hours after transfection with the pEGFP-SCdx plasmid showing a reduced fluorescence compared to similar timepoint without addition of retinoic acid (200x). **D)** HM-1 cells 12 hours after transfection with the pEGFP-ShortEco plasmid again showing no fluorescence above background (200x). **E)** The HM-1 cells transfected with pEGFP-LongEco showing lower levels of fluorescence than those seen in untreated cells (200x). **F)** HM-1 cells transfected with pEGFP-Sph again showing reduced levels of fluorescence (200x).

**Figure 4.18: Effect of retinoic acid on activity of Cdx1-EGFP reporter constructs in ES cells, 12 hours post-transfection**





## 4.6 Discussion

All homeobox genes code for proteins containing a cross-species conserved 61 amino acid helix-turn-helix homeobox domain, capable of binding to DNA and through this interaction influencing transcription from target genes. These homeobox genes are believed to be of central importance to the formation of body-axis during embryogenesis in almost all animal species. For this reason the understanding of the transcriptional control of homeobox genes could provide important clues to the signals involved in the complex process of tissue patterning in animal development. The method of transcriptional control of these genes may be of importance for other reasons, as well. More recently it has been shown that the regulation, or dysregulation, of certain homeobox genes may be of importance in vertebrate animals beyond embryogenesis. A number of homeobox genes have been found to demonstrate altered expression in neoplastic tissues of both the human and the mouse, including kidney, breast and colorectal lesions (Cillo *et al.*, 1992; Jin *et al.*, 1999; Vider *et al.*, 2000). Altered expression of several homeobox genes of the Hox family have also been found in human leukaemias and lymphomas, indeed these genes appear to possess oncogenic potential when overexpressed (Maulbecker and Gruss, 1993).

More specifically, the deregulation of the murine and human versions of *Cdx-1* has been implicated in tumourigenesis. Overexpression and nuclear localisation of *Cdx-1* has been observed in epithelial cells in intestinal metaplasia of the human stomach and oesophagus while *Cdx-1* levels appear to be significantly depressed in neoplastic regions of the large intestine (Silberg *et al.*, 1997). Co-transfection of the overexpression vectors for *Cdx-1* and a related homeobox gene, *Cdx-2*, into the colorectal derived cell line HT 29, shown to be deficient in expression of these two genes, has demonstrated that their reintroduction potentiates apoptosis and lowers motility and tumour formation in these cells (Mallo *et al.*, 1998).

The suppression of *Cdx-1* expression in neoplastic regions of the large intestine is very interesting given the recent demonstration that *Cdx-1* is a target of the Wnt/ $\beta$ -catenin/TCF signalling pathway (Lickert *et al.*, 2000). As previously

mentioned the large majority of tumours arising in the colon carry mutations in APC or  $\beta$ -catenin; these mutations either interrupt the ability of APC to bind  $\beta$ -catenin and hence target the protein for degradation or result in  $\beta$ -catenin molecules which are inherently resistant to APC induced phosphorylation (Polakis, 1999). This mutation spectrum is reflected in the genotypes of cell lines derived from human colorectal tumours (Dihlmann *et al.*, 1997; Ilyas *et al.*, 1997). The result of these mutations is the accumulation of  $\beta$ -catenin within the cell, ultimately resulting in transcriptional activation of targets of the TCF/LEF class of transcription factors. The demonstration that *Cdx-1* is a TCF/LEF target gene coupled with the knowledge that accumulation of  $\beta$ -catenin and hence transactivation of TCF/LEF targets is an early event in colorectal tumourigenesis would suggest that *Cdx-1* expression should rise during the initial stages of transformation from normal tissue to neoplastic lesion. While increase in *Cdx-1* is observed in gastric and oesophageal lesions this is not the case in the large intestine. This suggests that some means of reducing or stopping expression from the *Cdx-1* locus must be able to override the influence of the increased TCF/LEF transcriptional potential.

One possibility is that the *Cdx-1* alleles suffer inactivation mutations early on in the process of tumourigenesis. While no comprehensive study of *Cdx-1* mutation frequency in colorectal lesions has been performed, mice heterozygous for a null mutation in the related *Cdx-2* gene are prone to intestinal tumours and *Cdx-2* mutations have been noted in a number of replication error positive (RER+) human colorectal tumours, although the frequency of mutation seems to indicate that *Cdx-2* mutation does not contribute significantly to tumourigenesis (Yagi *et al.*, 1999). *Cdx-1* heterozygous mice do not appear to show the same predisposition to colorectal tumourigenesis. The relative rarity of *Cdx-1* mutations in colorectal tumours suggests that some other means of *Cdx-1* downregulation must be occurring in early neoplastic lesions of the large intestine.

The other control elements identified in the *Cdx-1* promoter sequence may play some role in this downregulation. If *Cdx-1* expression is driven by the presence

of retinoic acid in normal cells, the loss of ability to respond to normal physiological levels of retinoic acid could account for the observed drop in *Cdx-1* levels. Interestingly, retinoic acid is essential for maintenance of epithelial cell differentiation and appears to increase differentiation in a number of cancer-derived cell lines, it has been used as a chemotherapeutic to treat some types of poorly differentiated cancers (Carter *et al.*, 1996; Hansen *et al.*, 2000). It is possible that the observed reduction in *Cdx-1* expression in early intestinal lesions is indicative of a loss of retinoic acid responsiveness possibly preceding some degree of dedifferentiation. The expression from the *Cdx-1*-EGFP reporter plasmids did not increase at the dose previously reported to give maximal upregulation in the mouse embryo and embryonalcarcinoma cells, despite the fact that three of the reporter constructs contain the described retinoic acid response motif. It is possible that HCT 116 cells and murine primary colonocytes do not express the required retinoic acid binding protein to allow interaction with the novel *Cdx-1* retinoic acid response element. Embryonic stem cells more closely resemble the embryonalcarcinoma cells in which the retinoic acid dependant upregulation of *Cdx-1* was initially observed and may be expected to behave in a similar manner in response to retinoic acid. Despite their resemblance to embryonalcarcinoma cells, the ES cells transfected with the *Cdx-1*-EGFP reporter constructs showed reduced, rather than enhanced, reporter expression upon treatment with 1 $\mu$ M retinoic acid. Again, it is possible that the protein responsible for the interaction of retinoic acid with the novel retinoic acid response element found in the *Cdx-1* promoter is not present in high levels within ES cells. Perhaps more importantly, the ES cells responded to the presence of retinoic acid by slowing their growth and differentiating to some degree; treated wells were much less confluent after 2 days of observations than were corresponding wells of ES cells which had not been exposed to retinoic acid. The ability to actively drive transcription from the *Cdx-1*-EGFP reporter constructs may depend to some extent on whether the transfected cells are actively growing and dividing and so are susceptible to growth inhibition caused by retinoic acid exposure. Alternatively, *Cdx-1* expression may be downregulated during the process of differentiation, so that the exposure to retinoic acid causes a reduction in *Cdx-1* expression by driving this process. The observed pattern of *Cdx-1*

expression in the crypts of the murine intestine shows highest levels in the stem cells and proliferative compartments, with expression falling away as cells move out of this zone (Subramanian *et al.*, 1998). It is possible that this observed gradient of *Cdx-1* expression is related to the differentiation state of the cells as they migrate up the crypt wall and become progressively more differentiated.

High levels of expression are measured from the pEGFP-SCdx plasmid transfected cells; this plasmid contains 2 TCF/LEF sites and the positive epithelial element but does not contain the previously described silencer element. Transfection with the pEGFP-ShortEco plasmid produced a lower level of fluorescence intensity despite containing one more TCF/LEF binding site than the pEGFP-SCdx plasmid. This may be due to the presence of the silencer element within the pEGFP-ShortEco reporter construct, although this reporter also contains the described silencer blocking element. The level of fluorescence intensity in ES cells transfected with pEGFP-LongEco was higher than that measured in the pEGFP-ShortEco transfected cells, while the pEGFP-Sph transfected cells showed fluorescence intensity nearly equal that shown in the pEGFP-SCdx transfections. The increasing amount of promoter sequence in these reporters, perhaps most importantly the inclusion of additional TCF/LEF binding sites, appears to overcome the transcription blocking effects of the silencer element.

The differences in the expression levels observed between the various cell types may highlight known mechanisms of *Cdx-1* expression control, and in some cases points to the involvement of other regulatory processes. The lack of expression seen in the SW 480 cell line when compared to the HCT 116 cell line has a number of possible explanations. Perhaps the most likely is simply the reduced transfection efficiency seen in the SW 480 experiments, compounded by the relatively high autofluorescent background caused by their rounded morphology has created a situation where expression from the *Cdx-1*-EGFP constructs is masked in the relatively few cells which have been successfully transfected. Had the transfections of SW 480 cells proved as efficient and as easy to observe as those involving HCT 116 cells, some differences in the levels of expression from the *Cdx-1* promoter in these two cell lines would have been

expected.  $\beta$ -catenin exists in a number of distinct pools within cells; some of the protein is sequestered in its interactions with the membrane bound E-cadherin molecule, some  $\beta$ -catenin will be present in the nucleus in association with transcription factors and there also exists a pool of  $\beta$ -catenin associated with  $\alpha$ -catenin. The distribution of  $\beta$ -catenin between these pools appears to vary from cell line to cell line and distinct differences between the proportions of  $\beta$ -catenin in the various "pools" have been found in the colorectal carcinoma derived SW 480 cell line and the adenocarcinoma derived HCT 116 cell line when compared to MDCK cells.  $\beta$ -catenin in HCT 116 and SW 480 cells, in addition to being present in much greater concentrations, appears to be found in a larger proportion in a low molecular weight fraction; this appears to indicate a shift in the amount of  $\alpha$ -catenin associated  $\beta$ -catenin relative to the fraction bound to E-cadherin or nuclear pools (Stewart and Nelson, 1997). The relative amount of  $\beta$ -catenin represented in each of these pools differs between HCT 116 and SW 480 cells; HCT 116 cells contain a larger proportion of E-cadherin bound  $\beta$ -catenin. This difference may reflect a variation in the fraction of  $\beta$ -catenin which is free to bind TCF/LEF transcription factors and hence participate in the upregulation of Wnt target genes, such as *Cdx-1*. If high transfection efficiency had been achieved in SW 480 cells, it is possible that higher expression levels would have been observed. In HCT 116 cells significant expression was seen in cells transfected with the shortest reporter construct, pEGFP-SCdx, and from those transfected with the longest construct, pEGFP-Sph. The two intermediate reporter construct, pEGFP-ShortEco and pEGFP-LongEco, showed very little expression in HCT 116 cells. Expression from the pEGFP-SCdx reporter may be being driven by the previously described positive epithelial element present in the short stretch of the *Cdx-1* promoter built into this reporter. Additionally, this reporter construct does not contain the endogenous *Cdx-1* silencer (or silencer blocking element). The lack of transcription from pEGFP-ShortEco and pEGFP-LongEco in HCT 116 cells may be caused by the presence of this endogenous *Cdx-1* silencer element within the two intermediate reporter constructs; the silencer blocking element may not be able to completely overcome the transcriptional downregulating effect of the silencer without the fifth TCF/LEF



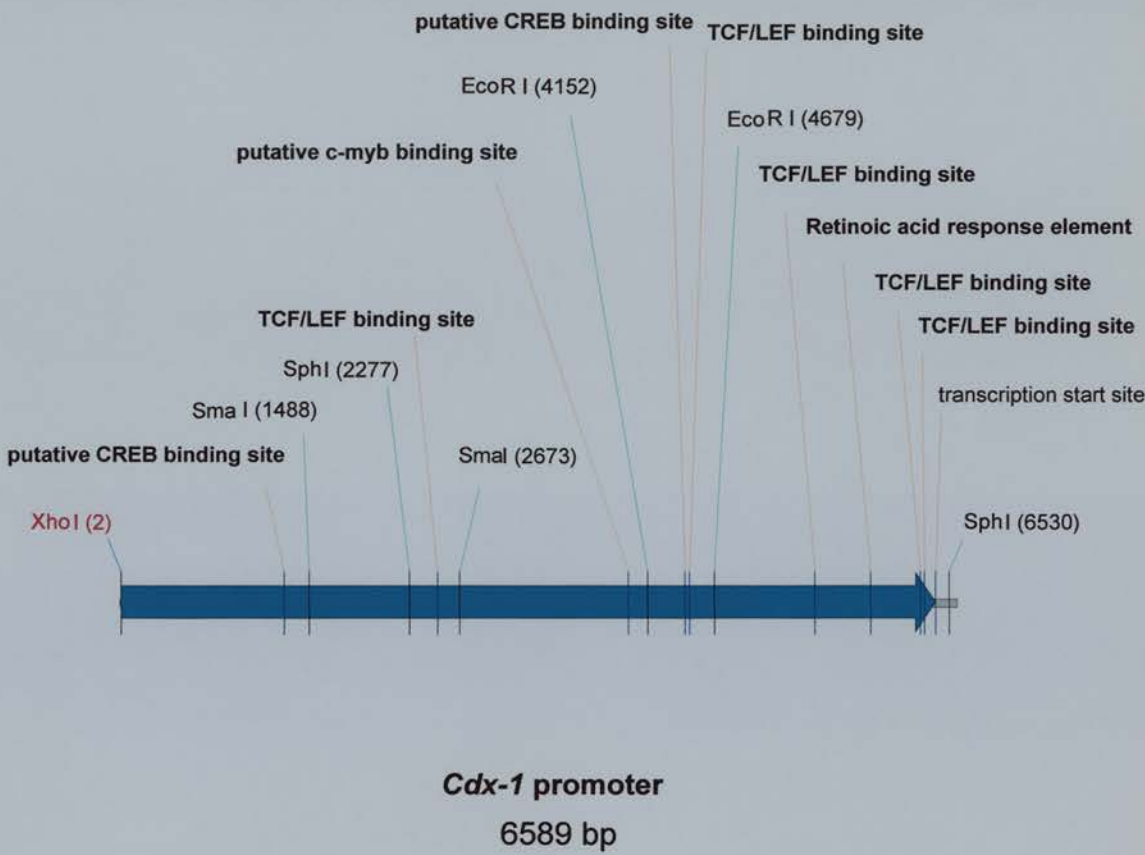
site or possibly other transcriptional elements found in the longer pEGFP-Sph reporter construct.

The *Cdx-1*-EGFP reporter containing the longest stretch of *Cdx-1* promoter, pEGFP-Sph, provokes much higher levels of EGFP expression in HCT 116 cells when compared to the pEGFP-LongEco plasmid. As previously mentioned, the pEGFP-Sph promoter region contains a fifth TCF/LEF binding site and as such could be expected to drive the highest levels of transcription as TCF/LEF binding is important in *Cdx-1* expression control. Interestingly, in murine HM-1 ES cells the highest levels of transcription are found in cells transfected with the pEGFP-LongEco plasmid, despite the fact that this plasmid contains one fewer TCF/LEF sites than pEGFP-Sph. An algorithm based search of the newly determined upstream *Cdx-1* sequence reveals a number of putative transcription factor binding sites which may be responsible for this observed difference in expression between murine ES cells and the HCT 116 human cell line (MatInspector, [www.gsf.de/cgi-bin/matsearch.pl](http://www.gsf.de/cgi-bin/matsearch.pl)). Perhaps most interesting is the presence of a c-myb binding site at -2426 relative to the transcription start site (see figure 4.19). This transcription factor is recognised as an oncogene and has been found to be upregulated in a number of different cancer and tumour derived cell lines, including colorectal cancers and colorectal cell lines (Ramsay *et al.*, 1992; Alitalo *et al.*, 1984). C-myb controls the expression of a number of genes important to cell cycle, apoptosis and also plays a role in the control of cell differentiation (Ramsay *et al.*, 2000; Thompson *et al.*, 1998; Del Bufalo *et al.*, 1996). ES cells express only low levels of c-myb, while HCT 116 cells may, in common with many colorectal derived cell lines examined to date, express very high levels of this transcription factor (Dyson *et al.*, 1989; Alitalo *et al.*, 1984; Untawale and Blick, 1988). High cellular levels of c-myb would be a potent factor in driving the expression of genes responsive to c-myb signalling, of which *Cdx-1* is a candidate, and lead to higher expression from the reporter construct containing the putative c-myb binding site in overexpressing cell lines. This may explain the higher relative level of EGFP transcription seen in the HCT 116 cell line with the pEGFP-Sph reporter compared with expression levels from the same reporter construct in HM-1 ES cells.



**Figure 4.19:** Diagram showing the locations of the characterised transcription factor binding sites in the murine *Cdx-1* promoter region in relation to some restriction enzyme sites used in the construction of the *Cdx-1*-EGFP reporter vectors. Also indicated are possible c-myb and CREB transcription factor binding sites identified by algorithm searches.

**Figure 4.19: Known and possible transcription factor binding sites in the murine *Cdx-1* promoter**



The low levels of expression from the *Cdx1*-EGFP reporters observed in the primary colonocyte preparations might again be the result of poor transfection efficiency coupled with high background masking low levels of EGFP fluorescence. An additional factor may have been the age of the cultures when transfected; the colonocyte preparations were generally transfected after 3 days in culture, the stem cells and the cells of the proliferative compartment which would be expected to express *Cdx-1* may have begun to differentiate into more mature, non-*Cdx-1* expressing cells prior to transfection. There are also reports of downregulation of *Cdx-1* expression in colonic epithelium co-cultured with some fibroblast cell lines; the primary cell preparations contained a significant amount of fibroblast contamination which may have provoked a downregulation of *Cdx-1* expression in the cultured colonocytes (Duluc *et al.*, 1997).

Algorithm searches of the newly determined upstream *Cdx-1* sequence reveal other potentially interesting transcription factor binding motifs, including a number of putative cyclic-AMP responsive element binding (CREB) sites (see figure 4.19). CREB plays a role in transcriptional regulation of a number of genes important to cancer development and progression, as does its associated protein and transcription factor CREB-binding protein (CBP). Notably, CBP appears to play a role in inhibiting  $\beta$ -catenin/TCF/LEF transcriptional activation, the close proximity of a TCF/LEF binding site within the *Cdx-1* promoter sequence may indicate that CBP is capable of downregulating *Cdx-1* expression (Barker *et al.*, 2000). *Cdx-1* is not the only homeobox gene to be known to contain CREB sites within its promoter; the related mouse homeobox gene *Cdx-2/3* contains a CREB consensus site and the *Drosophila* cardiac homeobox gene *tinman* depends on CREB sites within its promoter region to direct correct tissue-specific expression (Lorentz *et al.*, 1999; Venkatesh *et al.*, 2000). Interestingly, expression of the *tinman* homeobox gene is known to be modulated by *wingless*, the *Drosophila* Wnt-1 homologue, suggesting another similarity with murine *Cdx-1* (Wu *et al.*, 1995; Park *et al.*, 1996).

# **Chapter 5**

## **Characterisation of a transgenic floxed *Apc* mouse line**

## Chapter 5 - Characterisation of a transgenic Floxed *APC* mouse line

### 5.1 Background

As outlined in the introductory chapter, the *APC* gene and its protein product play a central role in the regulation of the Wnt-1 signalling pathway, a pathway whose dysregulation is heavily implicated in the formation and progression of colorectal tumours. The prevalence of *APC* mutations in spontaneous colorectal tumours, as well as the phenotype displayed by human patients with germline *APC* mutations, have pointed to *APC* mutations as early events in colorectal neoplasia and work with the *Min* mouse and other genetically altered *Apc* deficient strains have borne out this hypothesis. While these germline *Apc* truncation strains have many features mimicking the rare human condition FAP they are less accurate representations of the much more common phenomenon of spontaneously arising colorectal tumours. Ideally an animal model in which the *Apc* gene can be “switched off” at a precise time in the location most relevant to successful mimicking of the human spontaneous tumour, the stem cells of the large intestinal crypt, would provide a greater insight into the earliest cellular events in the progression from normal to neoplastic behaviour. The development of a large intestinal stem cell specific inducible Cre mouse strain, coupled with the production of a double floxed or floxed/null *Apc* mouse strain would allow this level of control. As mentioned in section 3.1, the ability to generate the *Apc* null genotype in the stem cell would also aid in the determination of whether *Apc* mutation alone is sufficient to provoke polyp (and eventually tumour) formation in the intestine. A project has been undertaken in Edinburgh to generate a transgenic mouse line bearing *APC* alleles whose entire lengths were flanked by loxP sites. This transgenic *CC* mouse line was generated through the work of a previous PhD student, Cindy Cooper; the analysis of the transgenic animals is presented in this chapter.

## 5.2 Derivation of floxed APC line

The transgene used by Ms. Cooper in the generation of the *CC* mouse line was assembled around a human cDNA clone supplied by Dr. Joanna Groden, University of Cincinnati, OH (Groden *et al.*, 1995). The *APC* cDNA was excised from its original plasmid backbone and through a number of cloning steps introduced into a new construct between two directly oriented loxP sites. The *APC* cDNA transgene expression construct was completed with the addition of the 3-phosphoglycerate kinase-1 (PGK) promoter and the SV40 polyadenylation sequences; the final transgene construct was termed pAPC 2M35 (see figure 5.1) (Adra *et al.*, 1987; Cooper, 2000). The PGK promoter is known to promote widespread, though tissue and cell type variable, expression in the mouse throughout development, as it is a necessary metabolic enzyme in the ubiquitous glycolytic pathway (McBurney *et al.*, 1994). Prior to use in pronuclear injections, Ms. Cooper successfully tested the transgene construct for its ability to undergo Cre mediated recombination reactions in a prokaryotic system, the constitutive Cre expression *E. Coli* strain BNN132 (Mo Bi Tech, GmbH). Introduction of the pAPC 2M35 plasmid into the BNN132 cells resulted in the appearance of DNA molecules consistent with the expected APC excision event (Cooper, 2000).

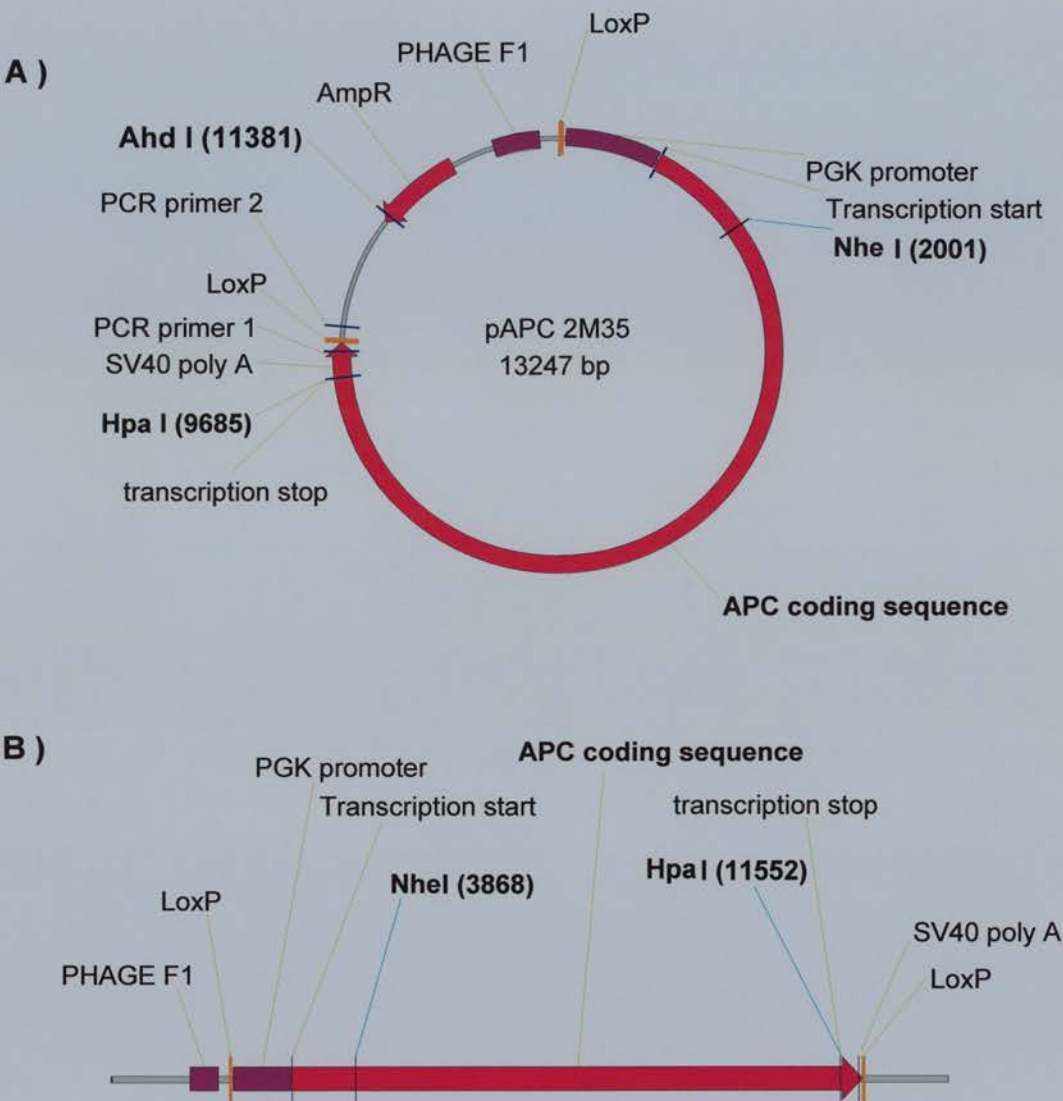
The pAPC 2M35 transgene was linearised with the restriction enzyme Ahd-I (see figure 5.1) and injected into eggs derived from F1 (C57BL/6 x CBA) x F1 (C57BL/6 x CBA) matings. One successful founder mouse was generated, proving positive for the presence of the pAPC 2M35 transgene by PCR reaction (see figure 5.3 for an example of the transgene detection PCR product). The founder mouse was immediately bred onto a C57BL/6 mouse line bearing the *Min* mutation in an effort to generate partial or complete rescue of *Min* homozygous mice which are normally embryonic lethal (Moser *et al.*, 1995). Detection of the transgenic sequence in early mice (transgene heterozygotes) by southern blot was not successful (Cooper, 2000). *CC* mice have also been screened for transgene expression using reverse transcriptase (RT) PCR and transgene expression was detected in several tissues (kidney, spleen, pancreas, muscle, large and small intestine, lung, heart, testis, liver, brain, thymus, mammary and salivary glands), however no information regarding the level of



expression could be inferred from the RT-PCR results (Cooper, 2000). Additionally, no information regarding expression from the pAPC 2M35 transgene could be determined by immunoblotting as the mouse and human proteins are similar enough to cause cross reactivity with any of the currently available antibodies directed against APC.

To determine if the pAPC 2M35 transgene was capable of undergoing Cre mediated recombination following its introduction into the mouse genome, Ms. Cooper derived embryonic fibroblasts from transgene positive embryos and infected them with an altered adenovirus coding for the Cre recombinase. Following infection with the adenovirus vector, DNA was extracted from the cells and screened by PCR for the presence of recombination products. The PCR used primers outside the floxed region of the pAPC 2M35 linearised transgene; recombination events would result in the primers spanning a region of 190bp while lack of recombination would leave the primer binding sites over 13kb apart, making PCR amplification unlikely. The detection of a 190bp band in infected transgene fibroblast populations indicated that Cre mediated recombination was successful (Cooper, 2000).

**Figure 5.1: pAPC 2M35 transgene vector**

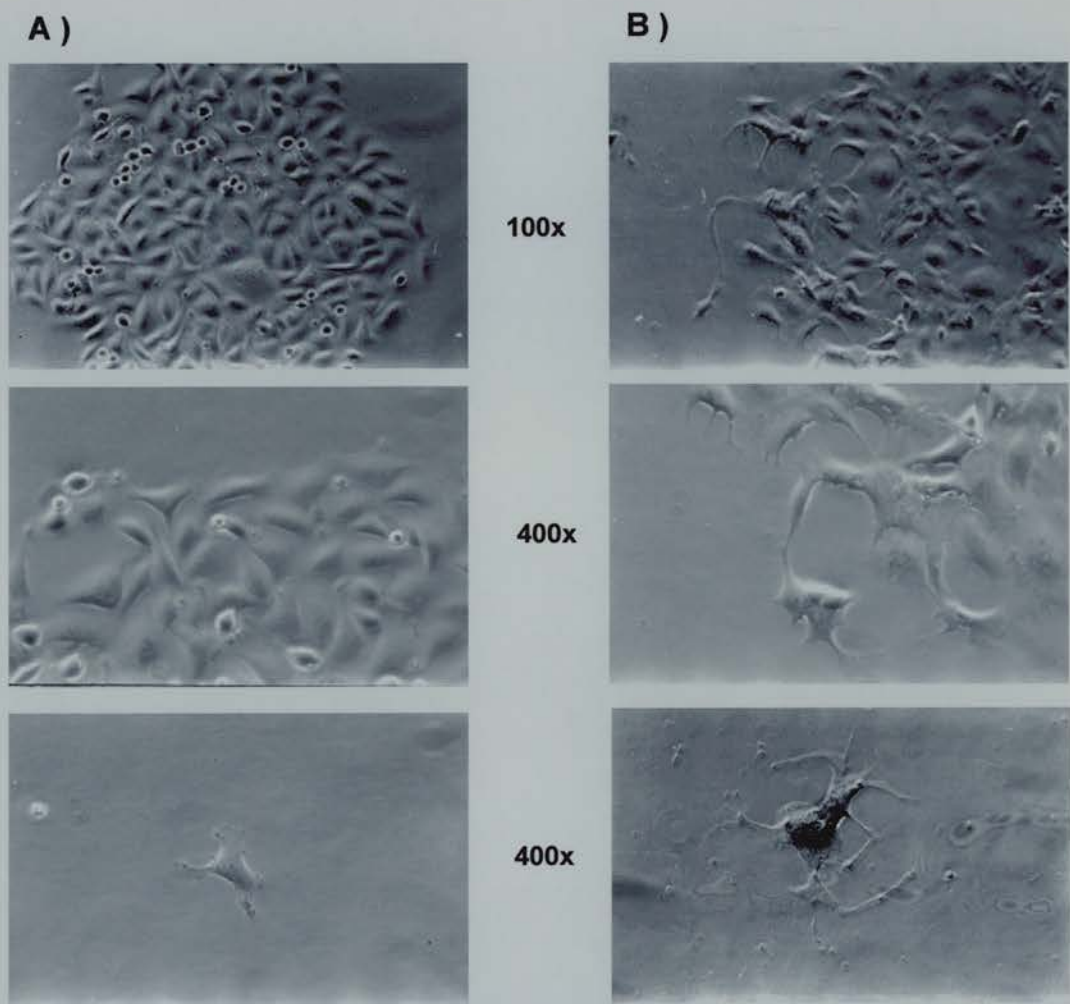


**Figure 5.1:** **A)** The floxed APC expression transgene vector constructed by Cindy Cooper. Expression of the human APC cDNA is driven by the mouse 3-phosphoglycerate kinase-1 (PGK-1) promoter. **B)** The pAPC 2M35 vector was linearised by digestion with the restriction enzyme AhdI prior to pronuclear injection.

### 5.2.1 *in vitro* analysis of transgene expression

In collaboration with Ms. Cooper, the Ahd-1 linearised construct was also introduced by calcium phosphate transfection into a number of human colorectal adenoma derived cell lines including DLD-1, HCT 116, HT 29 and SW 480. A linearised Neo resistance plasmid was co-transfected with the APC expression transgene and the cell lines were placed under G418 selection to isolate resistant cells. No stable clones were recovered from the cell lines DLD-1 or HT 29, which express a truncated version of the APC protein while one slow growing, morphologically altered clone (see figure 5.2) derived from SW 480, which also only expresses a truncated APC protein, survived for a brief period. Similar results were seen in another study involving the introduction of *APC* expression vectors into human colorectal cell lines (Grodén *et al.*, 1995). Several HCT 116 clones survived G418 selection and were shown by PCR (pAPC 2M35 transgene PCR, see section 2.2.2) to contain the APC transgene. The HCT 116 cell line normally expresses a full-length APC protein; a  $\beta$ -catenin mutation in this cell line hinders its phosphorylation, making it resistant to ubiquitination and subsequent degradation (Ilyas *et al.*, 1997). The transfections involving the pAPC 2M35 transgene suggested that this floxed human APC expression transgene was actively transcribed in these cell lines and that the translation product was capable of stopping (or greatly slowing) growth in those cells in which full length APC was normally absent.

**Figure 5.2:** Effect of introduction of *APC* transgene into SW 480 colorectal cell line

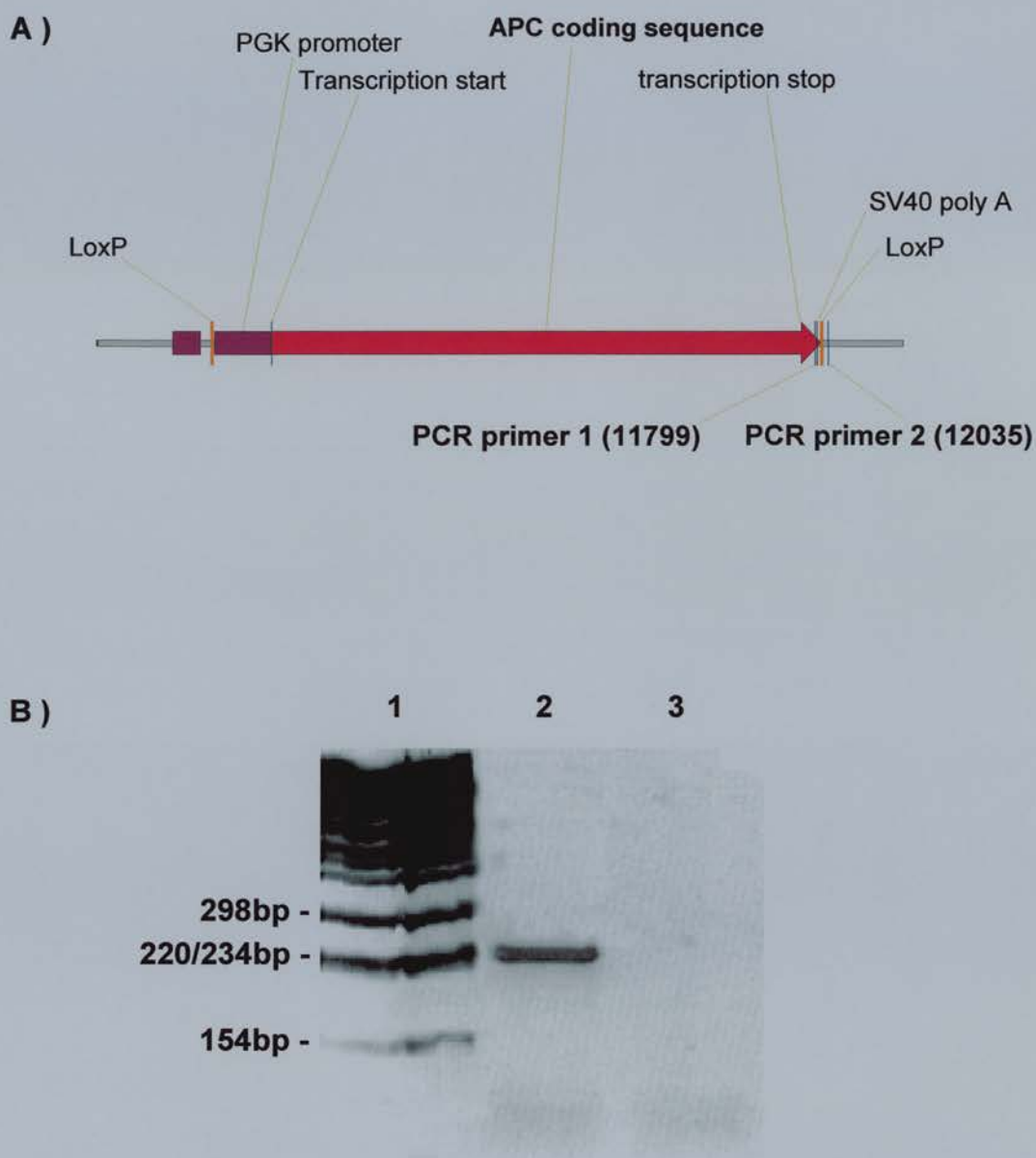


**Figure 5.2:** A) SW 480 cells which have been transfected with the *Neo/Tk* resistance vector alone showing normal morphology. B) Cells co-transfected with the *Neo/Tk* resistance vector and the pAPC 2M35 transgene vector showing an altered morphology. Magnification is indicated in the centre column.

### 5.3 Analysis of the CC transgenic mouse line

To date 207 informative progeny mice have been generated from the original founder mouse. CC mice derived from the original founder were screened for their *Min* status by PCR (see section 2.2.2) in addition to being screened for the presence of the pAPC 2M35 transgene (see section 2.2.2, figure 5.3); the pAPC 2M35 PCR does not distinguish between heterozygosity and homozygosity for the transgene. Three litters were excluded from the analysis of the CC mouse line. These litters were the product of matings between pAPC 2M35 heterozygous, *Apc*<sup>+/+</sup> animals and pAPC 2M35 negative, *Apc*<sup>Min/+</sup> animals. The litters containing 11, 6, and 9 mice (CC 26-51) were all weened in the same week and all animals proved to be transgene negative, *Apc*<sup>+/+</sup> on PCR analysis. The chance of generating this genotype from the described matings, assuming one site of transgene integration, is 0.25 for each offspring. Even for the smallest litter, the chance of the observed genotypes occurring by chance is 0.00024 (0.25<sup>6</sup>). Because the weenings all occurred in the same week it is assumed that these weening were mislabelled and they were excluded from all further analysis. The genotypes of mice generated to date from the various matings are summarised in the following table. Expected numbers were calculated assuming no embryonic lethality and a single transgene integration site.

**Figure 5.3: PCR detection of the integrated pAPC 2M35 transgene**



**Figure 5.3: A)** Vector diagram showing the expected state of the integrated pAPC 2M35 transgene; the PCR primers used to detect the sequence spanning the SV40 poly adenylation signal are indicated. **B)** Gel electrophoresis of the transgene sequence amplified from CC mice. Lane 1: Marker VI (Roche). Lane 2: PCR product amplified from genomic DNA from CC 240. Lane 3: The same amplification reaction performed using genomic DNA from CC 246.



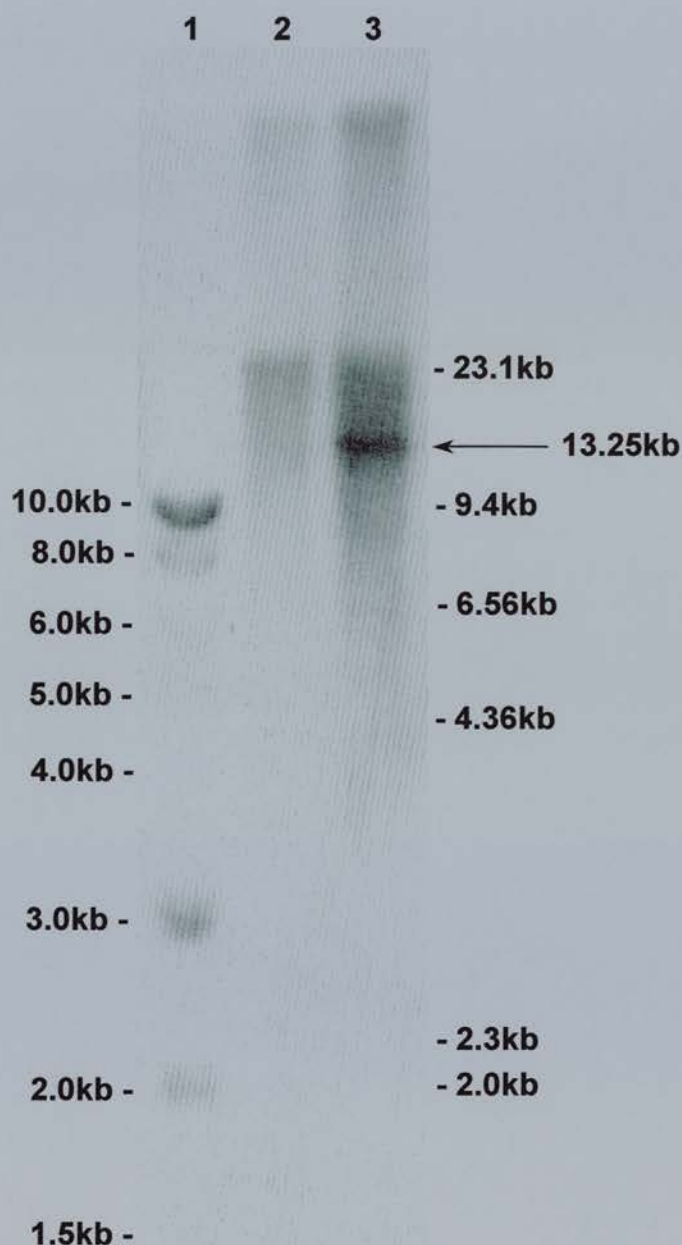
**Table 5.1: CC mouse line genotypes:**

Mating type	Possible outcome genotypes	Expected number of offspring of each genotype	Observed number of offspring of each genotype
<p><b>pAPC 2M35<sup>+/-</sup> Apc<sup>+/+</sup></b>  <b>x</b>  <b>pAPC 2M35<sup>-/-</sup> Apc<sup>Min/+</sup></b></p>	pAPC 2M35 <sup>+/-</sup> Apc <sup>+/+</sup>	16.5	15
	pAPC 2M35 <sup>+/-</sup> Apc <sup>Min/+</sup>	16.5	15
	pAPC 2M35 <sup>+/-</sup> Apc <sup>Min/Min</sup>	0	0
	Total pAPC 2M35 +	33	30
	pAPC 2M35 <sup>-/-</sup> Apc <sup>+/+</sup>	16.5	21
	pAPC 2M35 <sup>-/-</sup> Apc <sup>Min/+</sup>	16.5	15
	pAPC 2M35 <sup>-/-</sup> Apc <sup>Min/Min</sup>	0	0
	Total pAPC 2M35 -	33	36
	Total offspring	66	66
<p><b>pAPC 2M35<sup>+/-</sup> Apc<sup>Min/+</sup></b>  <b>x</b>  <b>pAPC 2M35<sup>-/-</sup> Apc<sup>Min/+</sup></b></p>	pAPC 2M35 <sup>+/-</sup> Apc <sup>+/+</sup>	4.125	9
	pAPC 2M35 <sup>+/-</sup> Apc <sup>Min/+</sup>	8.25	11
	pAPC 2M35 <sup>+/-</sup> Apc <sup>Min/Min</sup>	4.125	0
	Total pAPC 2M35 +	16.5	20
	pAPC 2M35 <sup>-/-</sup> Apc <sup>+/+</sup>	4.125	7
	pAPC 2M35 <sup>-/-</sup> Apc <sup>Min/+</sup>	8.25	6
	pAPC 2M35 <sup>-/-</sup> Apc <sup>Min/Min</sup>	4.125	0
	Total pAPC 2M35 -	16.5	13
	Total offspring	33	33
<p><b>pAPC 2M35<sup>+/-</sup> Apc<sup>Min/+</sup></b>  <b>x</b>  <b>pAPC 2M35<sup>+/-</sup> Apc<sup>Min/+</sup></b></p>	pAPC 2M35 <sup>+/?</sup> Apc <sup>+/+</sup>	20.25	21
	pAPC 2M35 <sup>+/?</sup> Apc <sup>Min/+</sup>	40.5	57
	pAPC 2M35 <sup>+/?</sup> Apc <sup>Min/Min</sup>	20.25	0
	Total pAPC 2M35 +	81	78
	pAPC 2M35 <sup>-/-</sup> Apc <sup>+/+</sup>	6.75	15
	pAPC 2M35 <sup>-/-</sup> Apc <sup>Min/+</sup>	13.5	15
	pAPC 2M35 <sup>-/-</sup> Apc <sup>Min/Min</sup>	6.75	0
	Total pAPC 2M35 -	27	30
	Total offspring	108	108
Total number of mice = 207			

Southern hybridisation analysis of DNA derived from later generations of *CC* transgenic mice (possible transgene homozygotes), digested with *NheI*, an enzyme which cuts once within the transgene sequence, has shown a band of the size expected from digestion of a tandem array of transgenes (see figure 5.4). To date no hybridisation bands corresponding to digestion of sequence flanking tandem array integration sites have been detected, preventing direct determination of the number of integration sites or approximate number of copies of the transgene present within the *CC* mouse genome by densitometry comparisons between the tandem array band and integration site bands. However, the number of integration sites for the pAPC 2M35 transgene can be estimated from the proportion of transgene positive mice generated from the mating data. The expected numbers of transgene positive offspring listed previously were generated assuming a single integration site of one transgene tandem array. With this assumption 130.5 transgene positive mice could be expected to be generated in the various matings. The actual number of transgene positive mice generated is 128; this ratio is statistically consistent with a single site of integration ( $\chi^2_1 = 0.041$ ,  $P = 0.839$ ).

Using the mating data it is also possible to determine that homozygosity for the pAPC 2M35 transgene is not embryonic lethal; matings between two transgene heterozygote parents produced 108 offspring. With a single site of tandem array integration 81 of these mice would be expected to be transgene positive, of which 33%, or 27 mice would be expected to be homozygous for the transgene tandem array. The observed number of transgene positive mice (heterozygotes and homozygotes) from these matings is 78. The observed number of transgene positive mice and the expected number of transgene positive mice are not significantly different ( $\chi^2_1 = 0.214$ ,  $P = 0.643$ ), suggesting that all transgene positive mice, regardless of transgene dosage, are surviving. Further, if transgene array homozygosity were lethal, then only 54 transgene positive mice would be expected to be generated out of 108 mice total. This is significantly different from the observed 78 transgene positive mice ( $\chi^2_1 = 11.221$ ,  $P = 0.001$ ), suggesting again that transgene array homozygosity is not lethal.

**Figure 5.4: Southern detection of pAPC 2M35 in transgenic mice**



**Figure 5.4: Southern detection of pAPC 2M35 in transgenic mice** Genomic DNA was prepared from liver samples obtained during dissections. The genomic DNA was digested overnight with a 10 fold excess of NheI which cuts once in the transgene sequence. The DNA blot was probed with a 236bp probe amplified by PCR from the SV40 poly adenylation sequence present in the transgene construct. Lane 1: Kilobase ladder (NEB). Lane 2: DNA from mouse CC 268, transgene negative by PCR screening. Lane 3: DNA from mouse CC 271, transgene positive by PCR. Not visible in the autoradiograph is Lane 4: Marker II (Roche) The expected 13.25kb transgene copy number band is clearly visible in the lane corresponding to the transgene positive DNA sample.

### 5.3.1 *Min* rescue

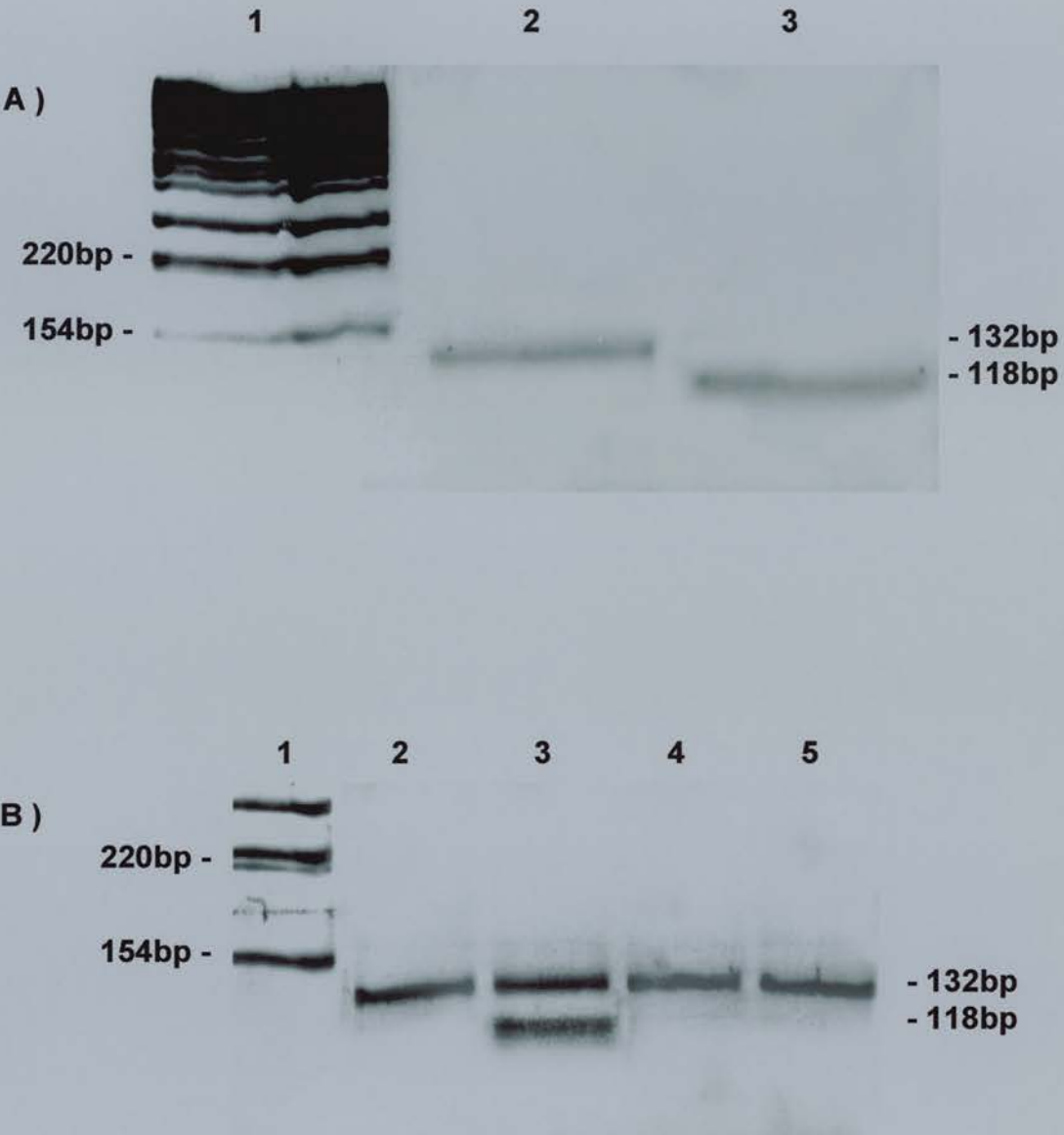
To date no *Min* homozygous mice have been generated; if transgene positivity (heterozygous or homozygous) were capable of rescuing the embryonic lethality of the *Min* homozygous phenotype the number of rescued mice would be expected to be approximately 25 (see table 5.1), significantly more than the number observed ( $\chi^2_1 = 27.432$ ,  $P < 0.001$ ). Additionally, approximately 7 (6.75) *Min* homozygotes could be expected out of the anticipated 27 transgene homozygous mice generated from the matings between transgene heterozygotes. The lack of *Min* rescue in the expected transgene homozygote population is also statistically significant ( $\chi^2_1 = 8.043$ ,  $P = 0.005$ ), indicating that inheritance of two transgene tandem arrays is not sufficient to overcome *Min* homozygous embryonic lethality.

Despite the lack of rescue, the APC expression transgene may have more subtle effects on the phenotype of the *Min* mice. One possibility is that the pAPC 2M35 transgene has some effect on the severity of the *Min* phenotype and as such affects survival times.

### 5.3.2 Comparison of survival times

The direct comparison of survival times within the CC colony is somewhat complicated by variation in the genetic background of the different generations of mice, as the background of a mouse bearing a *Min* mutation greatly affects the overall disease phenotype (Bilger *et al.*, 1996). The initial founder mouse possessed a 50% C57BL/6 / 50% CBA strain background; later generations are progressively more completely C57BL/6 (75% and 87.5%). In order to be able to compare survival times across generations, some means of correcting for the gradual change in background had to be found. While multiple genetic factors affect the *Min* phenotype, one particular gene polymorphism is known to contribute approximately 50% of the variation in tumour burden observed across different backgrounds (Dietrich *et al.*, 1993). The *Mom-1* locus (modifier of *Min-1*) has been identified as the secretory phospholipase-a2 (Pla2g2, see section 1.4.1); the C57BL/6 mouse strain bears a polymorphism of the *Mom-1* gene which results in high susceptibility to the *Min* phenotype (Gould *et al.*, 1996a; Gould *et al.*, 1996b). A number of polymorphisms of the *Mom-1* locus can be distinguished by PCR amplifications followed by digestion with BamHI (see section 2.2.2). The C57BL/6 inbred mouse strain gives a 138bp *Mom-1* genotyping PCR product which does not contain a BamHI site (Santos *et al.*, 1998). No published data regarding the size of the CBA *Mom-1* PCR/BamHI genotyping reaction could be found, although the *Mom-1* genotyping PCR products from a number of other inbred strains digest with BamHI to give shorter DNA molecules. It was determined that the *Mom-1* PCR product of the CBA strain is approximately 138bp in length and digests with BamHI to give a DNA molecule of approximately 118 bp (see figure 5.5). *Mom-1* PCR reactions and digests using both CBA and C57BL/6 genomic DNA as template showed that the polymorphisms could be successfully distinguished following electrophoresis on 4% agarose TBE gels (see chapter 2).

**Figure 5.5: Genotyping for *Mom-1* (Modifier of Min-1)**



**Figure 5.5:** **A)** *Mom-1* PCR genotype PCR product generated from a CBA strain genomic DNA sample. Lane 1: Marker VI (Roche), Lane 2: Undigested CBA *Mom-1* PCR product, 132bp. Lane 3: CBA *Mom-1* PCR product digested with BamHI, resulting in cleavage to a ~118bp product. **B)** Comparison of the CBA and C57BL/6 *Mom-1* PCR products. Lane 1: Marker VI (Roche). Lane 2: Undigested PCR product from mixed CBA / C57BL/6 genomic DNA. Lane 3: The mixed CBA / C57BL/6 PCR product digested with BamHI, showing the 132bp C57BL/6 and the ~118bp CBA bands. Lane 4: Undigested C57BL/6 *Mom-1* PCR product. Lane 5: C57BL/6 *Mom-1* PCR product digested with BamHI.

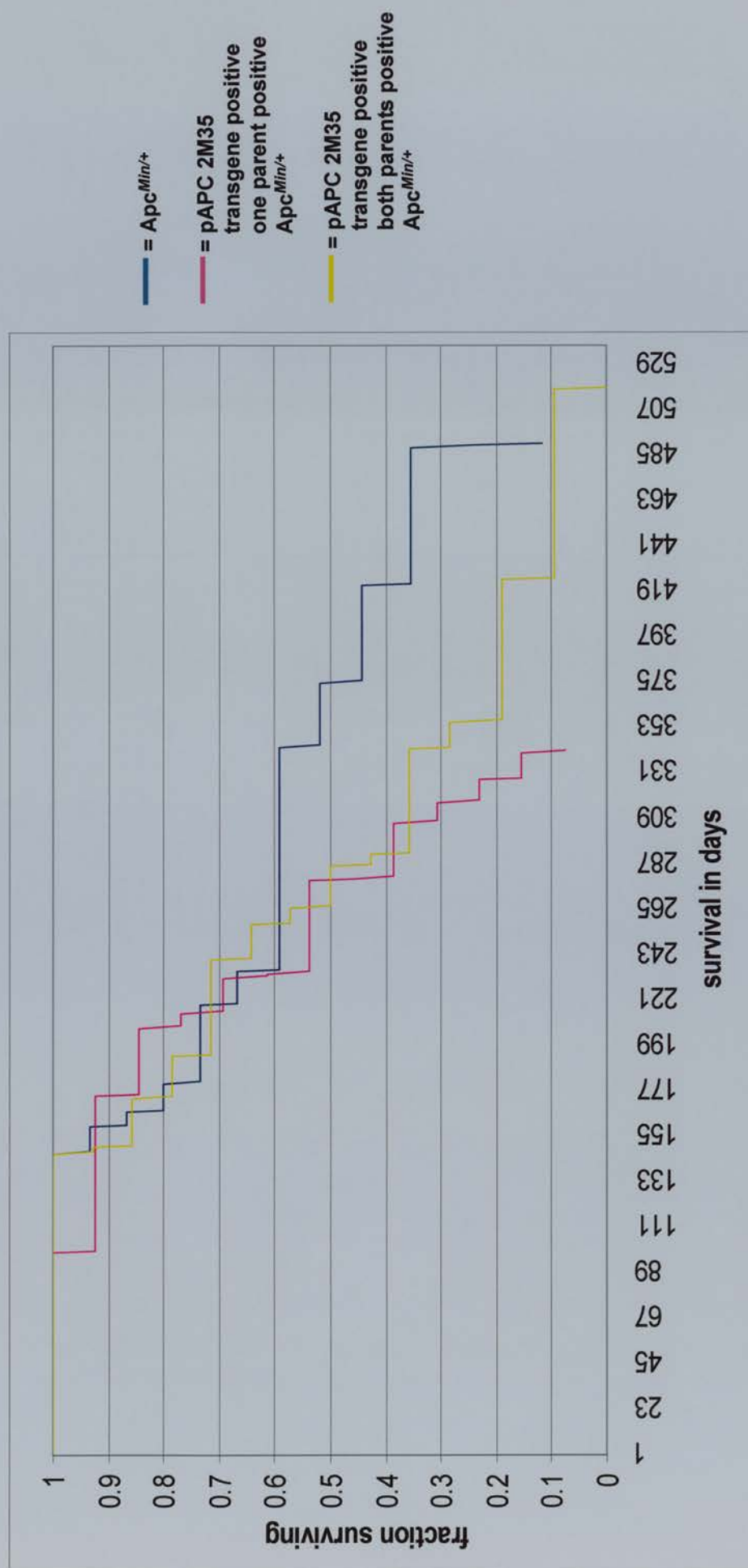


Genotyping genomic DNA samples from early CC mice determined that the founder mouse was homozygous for the C57BL/6 *Mom-1* polymorphism. As the *Min* mutant strain involved in subsequent matings is of a 100% C57BL/6 background, all CC mice are homozygous for this permissive C57BL/6 *Mom-1* polymorphism. While some variation may still result from the small inter-generation differences in strain background in the CC mice, the homogeneity of the *Mom-1* locus within the colony makes the direct comparison of survival times considerably more statistically valid.

Comparison of survival times involved 15 pAPC 2M35<sup>-/-</sup>, *Apc*<sup>Min/+</sup> mice, 13 pAPC 2M35<sup>+/-</sup>, *Apc*<sup>Min/+</sup> mice and 15 pAPC 2M35<sup>+/?</sup>, *Apc*<sup>Min/+</sup> mice, a proportion of which are expected to be transgene array homozygous. Survival of the transgene positive mice was analysed in groups based on the transgene status of their parents to determine if the mice which may have inherited transgene arrays from both parents had any difference in survival phenotype when compared to mice inheriting a single array of the pAPC 2M35 transgene. Statistical comparison of survival times was performed using the K-test (Lee, 1992). Survival times of mice which died of causes unrelated to tumour development (injuries or infections) or of unknown causes such as death during unmonitored periods when decay may obscure the precise cause of death were treated as censored data (see appendix 3 for raw data). Kaplan-Meier survival plots were also calculated and plotted (see figure 5.6) (Kaplan and Meier, 1958). The ranking of survival time used in the K-test indicate that no significant difference in the overall survival times exists between any of the groups of mice tested ( $\chi^2_1 = 2.96$ ,  $P > 0.05$ ).

**Figure 5.6:** The corrected survival curves for each mouse cohort. The dark blue curve represents the survival of transgene negative,  $Apc^{Min/+}$  mice. The mauve curve is represents the data from the transgene positive,  $Apc^{Min/+}$  mice with one transgene positive parent. The yellow curve is representative of the transgene positive,  $Apc^{Min/+}$  mice with two transgene positive parents.

Figure 5.6: Survival times of pAPC 2M35 transgene /  $Apc^{Min/+}$  and  $Apc^{Min/+}$  mice



5.3.3 Comparison of tumour burden

The consistent *Mom-1* background in the *CC* colony also allows for the comparison of tumour burden between transgene positive and transgene negative *Apc<sup>Min/+</sup>* mice. Haematoxylin and eosin stained sections of rolled gut preparations from 8 transgene positive and 8 transgene negative *Apc<sup>Min/+</sup>* mice were visually inspected for the presence of lesions (see figure 5.7). The intestinal sections were screened and the total number of lesions (small intestine and large intestine/colon) in each sample counted. The scoring was done without knowledge of the transgene status of the sample. While the number of tumours varied considerably from one individual animal to another, the median number of tumours per animal in each group (transgene positive and transgene negative) were very similar and a non-parametric test (Wilcoxon rank test) showed no significant difference between the two groups ( $P = 0.645$ ); the raw data are presented in the following table.

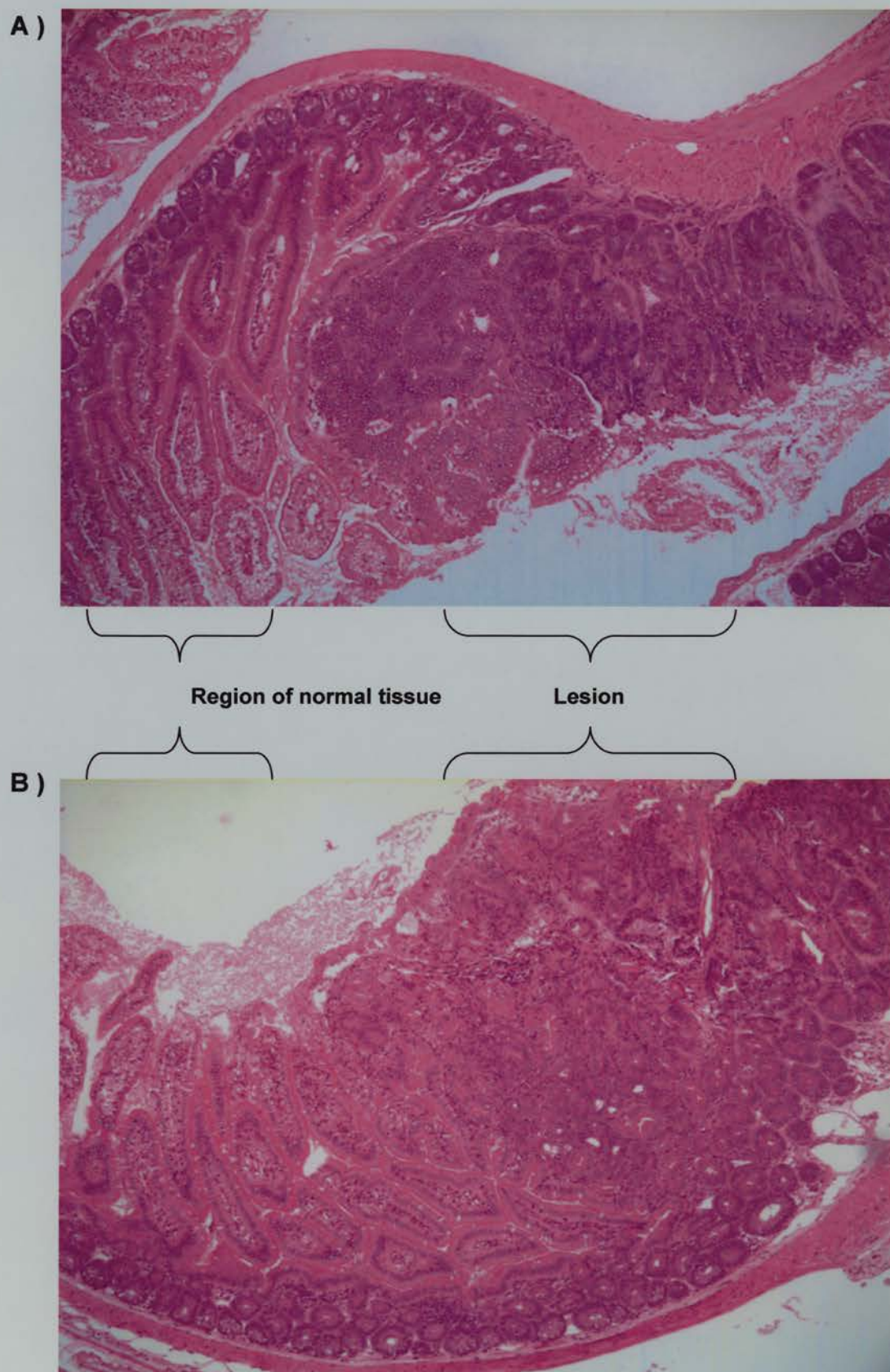
Table 5.2 Tumour burden data

Transgene positive mice			Transgene negative mice		
Tumour number			Tumour number		
Small intestine	Large intestine	Total	Small Intestine	Large intestine	Total
22	0	22	6	3	9
25	0	25	7	0	7
4	0	4	7	0	7
3	0	3	9	0	9
6	0	6	23	0	23
20	2	22	4	0	4
10	5	15	23	1	24
6	2	8	23	0	23

**Figure 5.7:** **A)** A large lesion of the small intestine in mouse CC 240, a transgene positive /  $Apc^{Min/+}$  animal. **B)** A similar lesion in the small intestine of mouse CC 246, a transgene negative /  $Apc^{Min/+}$  animal. Images are magnified ~200x.



**Figure 5.7: Small intestinal tumours in pAPC 2M35 transgene, *Apc<sup>Min/+</sup>* and *Apc<sup>Min/+</sup>* littermates**





### 5.3.4 Comparison of histological sections

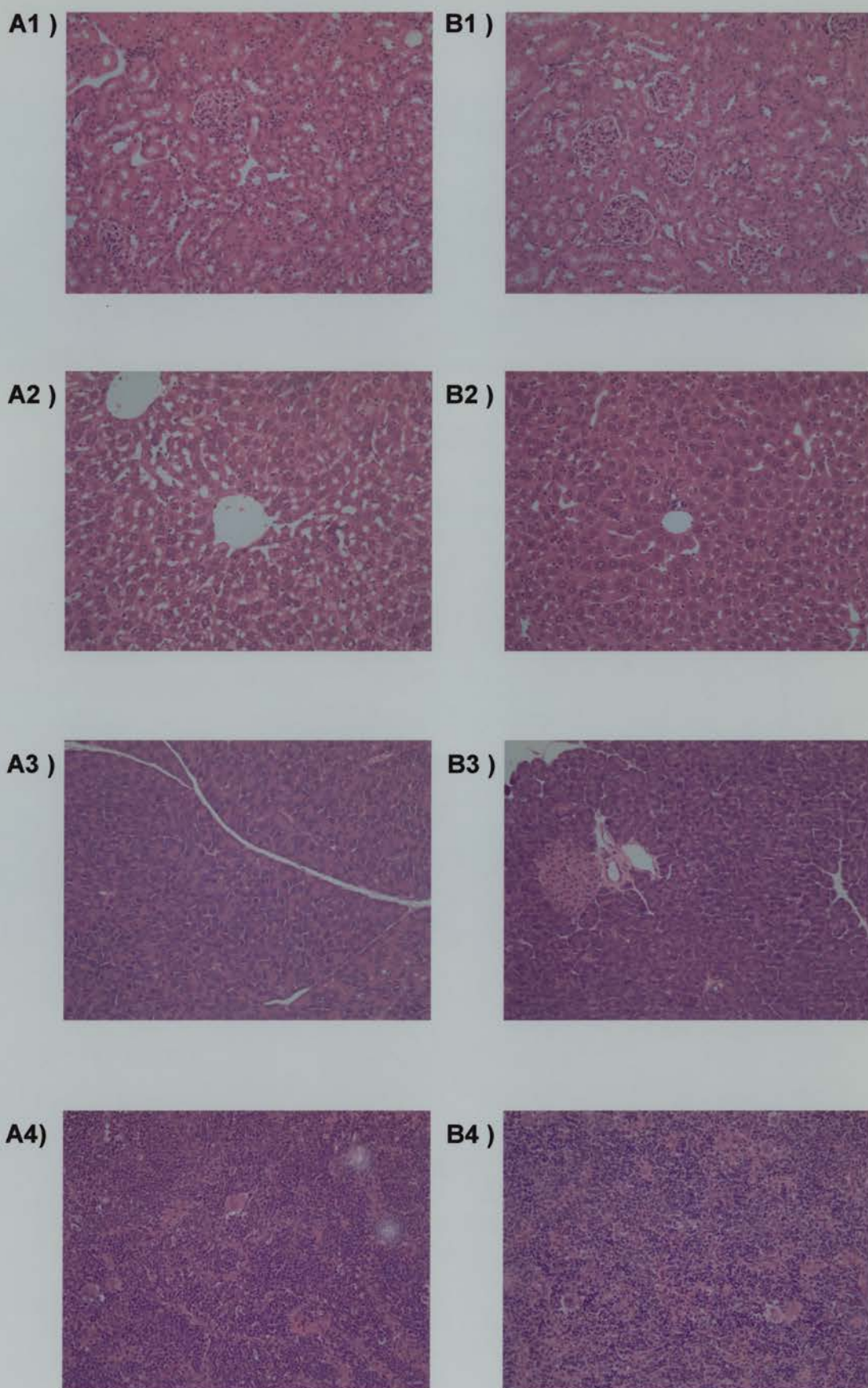
In addition to intestinal samples, a number of other tissue samples were taken during dissection of harvested CC mice. No consistent morphological abnormalities were observed in these tissues during dissection and this was confirmed by histological examination. Sectioned and haematoxylin/eosin stained samples of liver, pancreas, kidney and spleen from transgene positive and negative  $Apc^{Min/+}$  mice were examined for any histological differences (see figure 5.8). No detectable transgene associated changes could be discerned in the samples examined.

### 5.3.5 Transgenic embryos

As previously mentioned correct regulation of the Wnt-1 signalling pathway is essential during embryogenesis in a wide variety of species. The introduction of the pAPC 2M35 transgene may interfere with the developmentally regulated Wnt-1 pathway by provoking the expression of APC independent of developmental signals. A previous embryo screen involving transgene positive  $Apc^{Min/+}$  mice had found 2 embryos (day 9 and 10), which exhibited exencephaly (Cooper, 2000). In order to determine if the pAPC 2M35 transgene has an effect on the process of embryogenesis in the absence of the *Min* mutation, matings were set up between a transgene positive,  $Apc^{+/+}$  male and two transgene negative,  $Apc^{+/+}$  females. Pregnant females were harvested at day 8 to 10 of gestation and the embryos PCR screened for the presence of the pAPC 2M35 transgene (see section 2.2.2) and visually inspected under a dissection microscope for any developmental abnormalities (see figure 5.9). This set of timed matings generated a total of 15 embryos, 14 of which proved to be transgene positive. No developmental abnormalities were noted in any of the embryos, in particular there was no suggestion of failure of neural tube closure or malformation of the head.

**Figure 5.8:** Haematoxylin and eosin stained sections from four representative tissues. Images in the column “A” are from the transgene positive mouse CC 10. Images in the column “B” are from the transgene negative mouse CC 63. **A1/B1)** Kidney samples. **A2/B2)** Liver samples. **A3/B3)** Pancreas samples. **A4/B4)** Spleen samples. The transgene positive /  $Apc^{Min/+}$  mice showed no obvious histological alterations when compared with transgene negative /  $Apc^{Min/+}$  mice. All images are at ~200x magnification.

**Figure 5.8: Histology of pAPC 2M35 transgene /Apc<sup>Min/+</sup> and Apc<sup>Min/+</sup> mice**



**Figure 5.9:** pAPC 2M35 transgenic embryo



**Figure 5.9:** Example of a pAPC 2M35 transgene positive, *Apc*<sup>+/+</sup> embryo. The gross morphologies of the 14 transgene positive embryos were indistinguishable from similarly aged transgene negative embryos. 25x magnification.

## 5.4 Discussion

While the breeding and analysis of the *CC* line was underway, at least one other mouse strain bearing floxed *Apc* alleles has been developed (Shibata *et al.*, 1997). The homozygous *Apc*<sup>580S</sup> line appears phenotypically normal, displaying none of the phenotype associated with the *Min* mouse, despite the fact that the 14<sup>th</sup> exon of both *Apc* alleles has been flanked by loxP sites and the expression level of *Apc* from the altered loci appears to be only 30% that found in normal mice. On introduction of Cre-encoding adenovirus, via enema, these mice rapidly develop intestinal polyps, some of which progress to adenomas. The adenomas seen in these mice appear to be lacking the floxed exon 14, indicating that adenoma formation is correlated to successful recombination events. As previously mentioned, however, the adenovirus mediated excision in the *Apc*<sup>580S</sup> line fails to answer the question of whether loss of function of *Apc* alone is sufficient to provoke tumour formation. The development of further floxed *Apc* lines and the advent of an intestinal stem cell specific Cre line will better enable the investigation of the early event of colorectal carcinogenesis. The evidence collected to date, however, suggests that the presence of the pAPC 2M35 transgene has no statistically significant effect on the *Min* phenotype. In particular, no survival benefit (embryonic or adult) appears to be conferred on transgene positive animals, nor is there a significant reduction in the disease phenotype as evidenced by the tumour burden data.

There are a number of possible reasons for the lack of altered *Min* phenotype exhibited by pAPC 2M35 transgene *CC* mice. Although the original founder mouse was determined to harbour the transgene by PCR, this provides no information regarding the number of copies of the transgene present within the genome. Attempts to detect the transgene by Southern hybridisation in early mice failed, although faint signal was detected in genomic digestions spiked with 10 copies per mouse genome of the pAPC 2M35 transgene vector (Cindy Cooper, personal communication). While not strictly statistically valid, this suggested that the transgene was present in low numbers (<10) within the early *CC* mice. The detection of a transgene tandem array band by southern hybridisation indicates



that there exist at least 2 head-to-tail copies of the transgene inserted into at least one site.

While the PGK-1 promoter used in the pAPC 2M35 transgene is known to give global expression, the levels of expression can vary significantly from tissue to tissue (McBurney *et al.*, 1994). It is possible that the PGK-1 promoter does not promote sufficient expression in cells critical to embryonic survival to allow the PGK-1 driven expression of APC to compensate for lack of full length Apc protein in Apc<sup>Min/Min</sup> embryos, although PGK-1 expression appears to be upregulated during gastrulation, the stage of lethality for *Min* homozygotic embryos (Sturm *et al.*, 1999). PGK-1 activity also appears to be heavily influenced by methylation during development; it is possible that the PGK-1 promoter is downregulated in the same manner (Pravtcheva *et al.*, 1991).

Another possible influence on expression from the pAPC 2M35 transgene is the location of integration into the mouse genome. It is known that the integration of transgenic material into certain chromosomal regions can result in suppression of transgene expression, even in transgenes using very strong promoters. This may be due to transgenes integrating near endogenous silencer element or into methylation or architecturally silenced regions of chromosomes. The patterns of inheritance of the transgene in early CC mice suggests that the transgene had initially integrated into only one location; this leaves open the possibility that the transgene has simply been incorporated into the genome in such a way that it can drive only limited expression. Southern detection of the transgene was possible in later CC mice which may harbour copies of the transgene array inherited from both parents, however no information regarding the number of integration sites could be determined from the autoradiograph.

To date it has not been possible to quantify expression from the pAPC 2M35 transgene beyond determining by RT-PCR that transgene mRNA is being produced in a global manner. Immunoblot analysis would be inconclusive as the homology between the mouse and human protein is such that currently available antibodies would cross react between the endogenous mouse Apc and the human APC. Additionally, the similarity in the DNA sequence between the mouse and



human *APC* gene would make it very difficult to distinguish between mRNA transcribed from the endogenous mouse *Apc* locus and mRNAs transcribed from the transgene array using a northern blot technique.

While the pAPC 2M35 transgene is certainly present within the genome of the *CC* mouse strain and some expression can be detected, the lack of modulation of the *Min* phenotype makes the *CC* mouse line of limited use in the study of *Apc* dysregulation induced disease states.

# **Chapter 6**

## **Methylation and manipulation of the Wnt-1 pathway**

## Chapter 6- Methylation and manipulation of the Wnt-1 pathway

### 6.1 Background

Any cell which comes into contact with the outside environment or with the metabolic products of normal physiology will eventually encounter potentially genotoxic DNA damaging agents. These agents include reactive oxygen species generated during aerobic metabolism, radiation in the form of UV light and a host of chemical agents found in air, food and water. One of the most common types of genotoxic chemical agents encountered are alkylating agents. These reactive molecules are capable of bonding hydrocarbon groups to many types of normal cellular molecules including proteins and DNA. While the addition of these adducts to peptide macromolecules essential for cellular functioning can hinder or alter their normal spectrum of activity and may lead to cell death or apoptosis, the addition of these chemical groups to DNA can result in the generation of mutations if the adducts are present during DNA synthesis. In response to their constant exposure to alkylating agents all cells have evolved mechanisms to recognise and remove or replace nucleotides which have been altered by alkylation, or, if the damage is too great, to provoke apoptosis. These mechanisms appear to be important in the suppression of tumour formation; germline mutations in the *MSH2* gene, involved in the cellular recognition and response to a particular class of nucleotide methyl-group alterations, appear to cause some cases of the human HNPCC phenotype (see section 1.1.2.1) while many different types of cancers have been found to bear mutations in the p53 tumour suppressor, another protein involved in cellular responses to various forms of DNA damage.

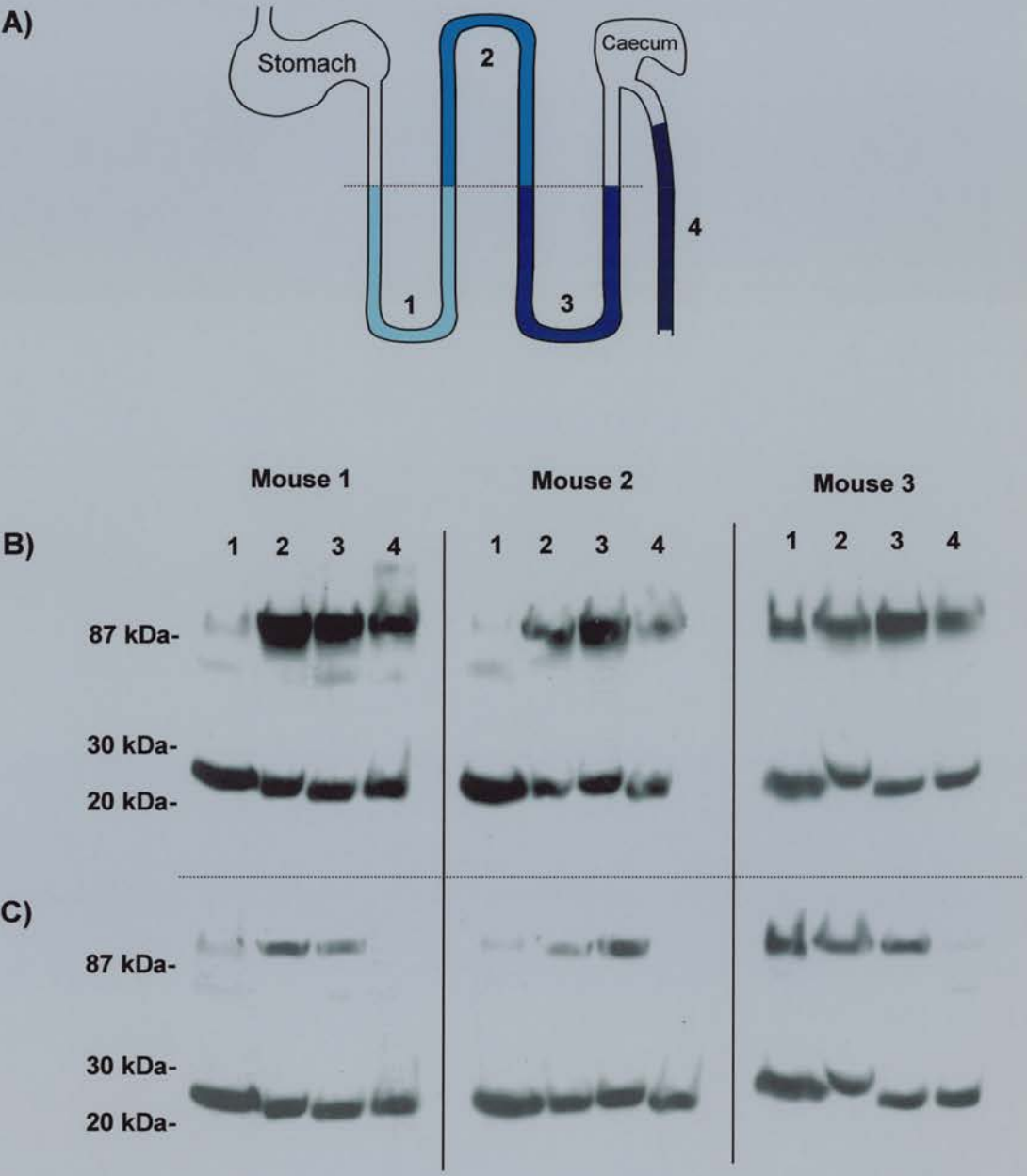
There appears to be a complex relationship between the *MSH2* mismatch repair (MMR) mechanisms of mammalian cells and the tumour suppressor protein p53. Mutations in either system appear to decrease the incidence of apoptosis, although the mechanisms by which each of these proteins influence apoptotic potential appear to be at least partially distinct (Clarke *et al.*, 1994; Dou *et al.*, 1995). While the two systems are able to initiate apoptosis in response to DNA damage it now appears that the apoptotic response to DNA alkylation is

primarily under the control of the MMR mechanisms alone (Toft *et al.*, 1999; Dou *et al.*, 1995). Despite this apparent independence it has been shown that the p53 pathway requires the presence of functional MMR mechanisms in order to respond to certain specific classes of DNA methylation damage; the MutS $\alpha$  pathway of MMR, involving heterodimers of *MSH2* and the related protein *MSH6*, is responsible for the stabilisation and subsequent accumulation of p53 in response to the presence of *O*<sup>6</sup>-methylguanine (*O*<sup>6</sup>MeG) DNA adducts (Hickman and Samson, 1999). The *O*<sup>6</sup>MeG DNA adduct is one of several genotoxic insults resulting from exposure to the powerful methylating agent *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (MNNG). As mentioned previously (see section 1.3.3) the levels of both the APC protein and its precursor mRNA are upregulated in a human colorectal cell line and in MEFs in response to MNNG exposure; additionally this upregulation has been shown to be dependant on the presence of functional p53 within the treated cell (Jaiswal and Narayan, 1998; Narayan and Jaiswal, 1997). The p53 dependant correlation between methylation damage and cellular levels of the APC tumour suppressor protein may have physiological significance; the levels of *O*<sup>6</sup>MeG damage on normal DNA in cancer prone regions of the human large intestine have been found to be elevated (Povey *et al.*, 2000). The ability for this type of adduct, frequently encountered in the relatively highly genotoxic environment of the large intestine, to upregulate cellular levels of the tumour suppressor most often mutated in colorectal tumours may represent an important protective mechanism. p53 dependant upregulation of APC in response to this type of methylation may help prevent neoplastic behaviour in *O*<sup>6</sup>MeG damaged cells by indirectly inducing downregulation of Wnt-1 responsive genes, many of which have been shown to have oncogenic potential (see section 1.3.1.6.2).

## 6.2 *in vivo* MNNG treatment

Initial experiments were carried out to determine whether this p53 dependant Apc upregulation in response to MNNG treatment could be observed *in vivo*. Immunodetection to compare cellular levels of the Apc target protein  $\beta$ -catenin in the various fibroblast populations would also allow observation of the immediate downstream Wnt-1 pathway effects of the expected Apc upregulation. Levels of  $\beta$ -catenin in protein extracts derived from murine intestine were compared between wild type mice treated with 50 $\mu$ g/kg MNNG and untreated controls, to determine if the expected reduction of  $\beta$ -catenin levels due to Apc upregulation could be detected. Unfortunately, the levels of  $\beta$ -catenin appeared to vary considerably along the length of the small intestine and there also proved to be considerable variation between protein samples prepared from individual mice (see figure 6.1). As this inter-sample variation would make subtle changes in  $\beta$ -catenin levels difficult to detect, further experiments were carried out using murine embryonic fibroblasts (MEFs). In an effort to determine whether the accumulation of Apc protein following MNNG exposure was also dependant on the presence of functional Msh2 protein, MEFs bearing the relevant p53 and/or *Msh2* homozygous null mutations were isolated and exposed to significant doses of MNNG.

**Figure 6.1:**  $\beta$ -catenin levels in the murine intestine



**Figure 6.1:** Western blot analysis of  $\beta$ -catenin levels in various sections of the murine intestine showing high variability. **A)** Diagram showing the dissection of the intestine for protein extraction. 1) proximal small intestine 2) mid-small intestine 3) distal small intestine 4) large intestine/colon **B)** 90kDa signal corresponding to  $\beta$ -catenin, and a 25 kDa signal corresponding to IgG light chain.  $\beta$ -catenin was detected with an antibody directed against an epitope coded for in exon 14. **C)** The same samples probed with an antibody directed against an epitope coded for in exon 2 of  $\beta$ -catenin.



## **6.3 Embryonic fibroblasts**

### **6.3.1 Preparation and genotyping**

Embryonic fibroblasts were generated using the protocol outlined in section 2.3.1. The embryos used in the generation of MEFs were derived from matings between pairs of animals heterozygous and/or homozygous for the null alleles of one or both of p53 and *Msh2*. In all cases a small tissue sample was taken at the time of embryo harvesting in order to PCR genotype the individual embryos for the relevant mutations (see section 2.2.2 for PCR conditions, see figure 6.2a and 6.2b). The cells were cultured, treated with MNNG and harvested for protein as outlined in section 2.3.

### **6.3.2 MNNG induced alteration of $\beta$ -catenin levels**

The upregulation of APC levels in MEFs observed by Jaiswal *et al.* appeared to be optimal after 15 hours of exposure to 50 $\mu$ M concentrations of MNNG in 0.5% serum. Accordingly, the various fibroblast populations in the present study were each split into 2 flasks, one being treated with a 50 $\mu$ M concentration of the methylating agent for 15 hours prior to protein harvesting and the other being an untreated control. The protein concentrations in the extracts were compared (see section 2.4.2) and the samples adjusted in order to ensure equal loading prior to gel electrophoresis; immunoblots were prepared as previously outlined and probed with monoclonal antibodies directed against epitopes coded for in either exon 14 or exon 2 of  $\beta$ -catenin (exon 14, C-terminal antibody: Clone 14, Transduction Laboratories; exon 2, N-terminal antibody: Clone 7D11, Upstate Biotechnology).

#### **6.3.2.1 $\beta$ -catenin levels in wild type fibroblasts**

Untreated wild type fibroblasts displayed a strong signal band at the expected size of ~90kDa after culture in 0.5% serum for 15 hours. On administration of MNNG the relative level of  $\beta$ -catenin signal observed appeared to drop significantly (see figure 6.2). Densitometry analysis of the bands on a sub-

saturation film exposure confirmed the initial observations of reduced  $\beta$ -catenin signal. Densitometry analysis was further corrected for gel loading by comparison of  $\beta$ -catenin signal with densitometry of an unrelated high molecular weight band in the Coomassie stained post-transfer gel (see figure 6.2b and 6.2c). Corrected densitometry data confirmed the reduction of  $\beta$ -catenin signal following MNNG treatment.

#### **6.3.2.2 $\beta$ -catenin levels in $p53^{-/-}$ fibroblasts**

An identical comparison of the  $\beta$ -catenin levels in treated and untreated  $p53$  null fibroblasts revealed a small reduction in  $\beta$ -catenin signal strength, which was related to MNNG dose (see figure 6.3). The reduction was far less distinct than that observed in the wild type fibroblast preparations; further loading correction performed as outlined above proved the slight observed reduction to be even smaller than initially measured.

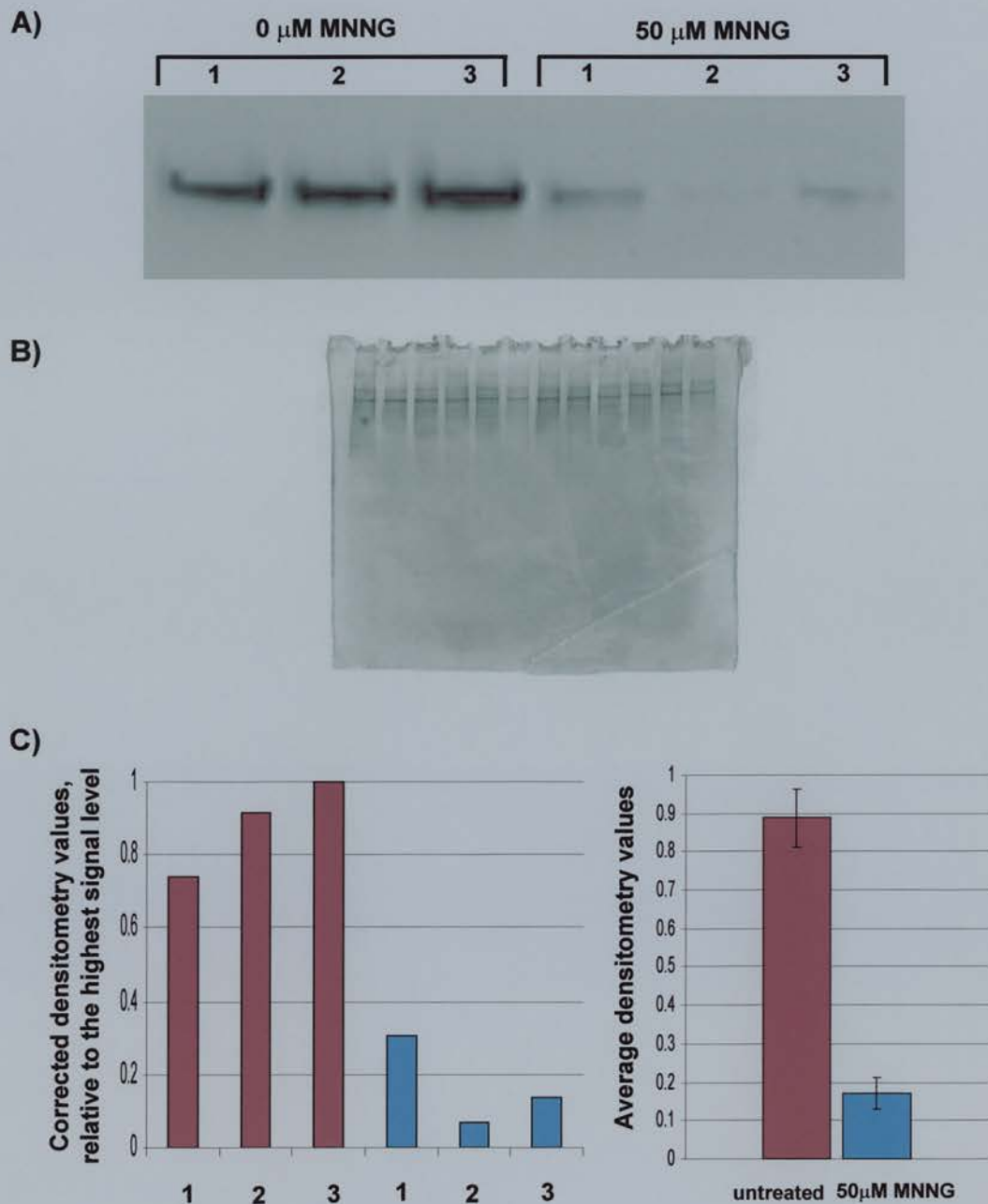
#### **6.3.2.3 $\beta$ -catenin levels in $Msh2^{-/-}$ fibroblasts**

*Msh2* null fibroblasts, which retained wild type alleles of  $p53$ , showed a significant drop in  $\beta$ -catenin levels following MNNG exposure (see figure 6.4). Loading correction performed as above confirmed this observed  $\beta$ -catenin reduction and showed that the level of reduction was comparable to that observed in the wild type fibroblast experiments.

#### **6.3.2.4 $\beta$ -catenin levels in $p53^{-/-}/Msh2^{-/-}$ fibroblasts**

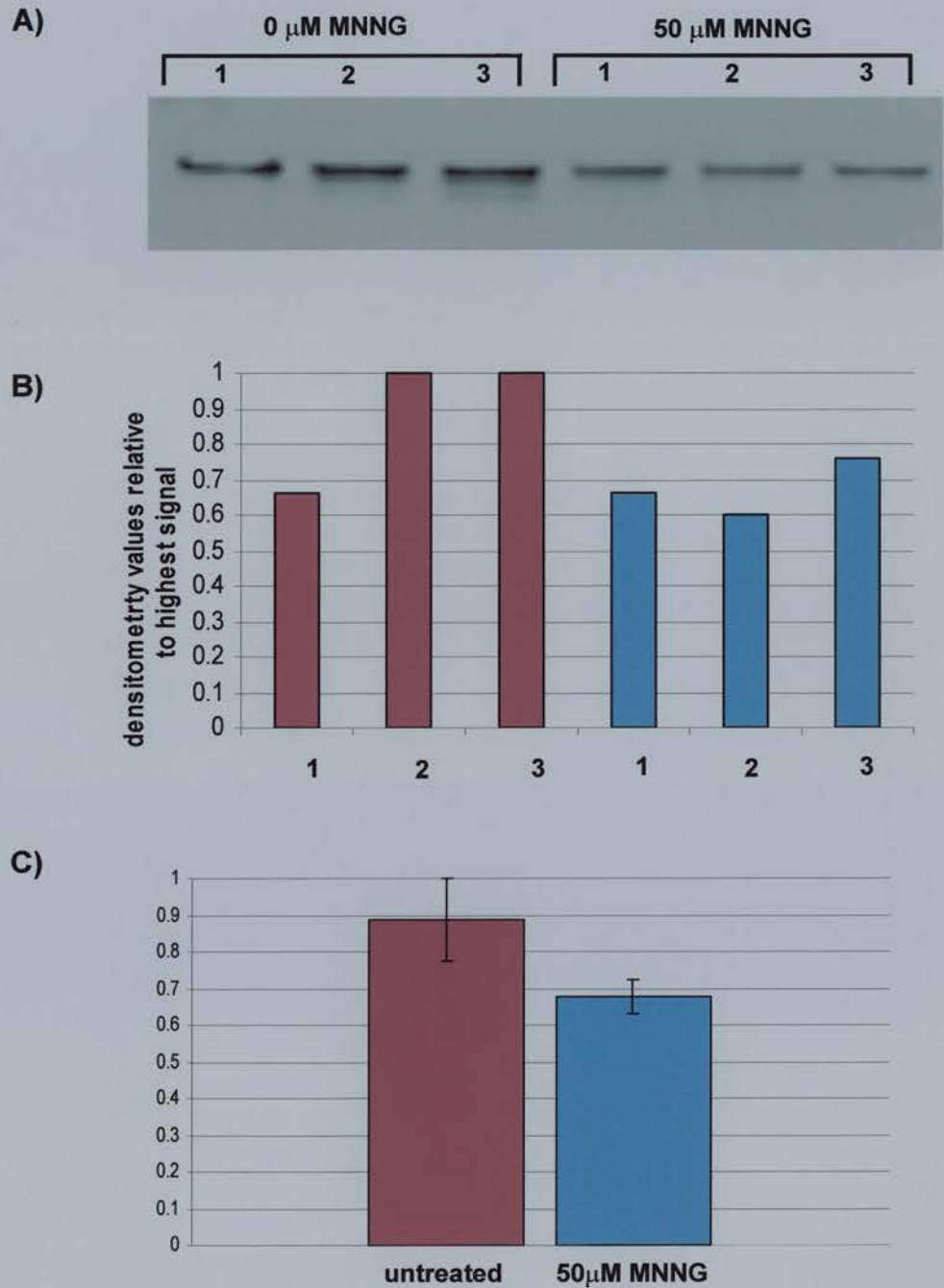
Double null fibroblasts, bearing homozygous mutations in both  $p53$  and *Msh2* alleles were also subjected to MNNG treatment. Analysis of  $\beta$ -catenin levels before and after treatment showed a small level of  $\beta$ -catenin reduction, comparable to that observed in the  $p53$  single null cells (see figure 6.5).

**Figure 6.2:**  $\beta$ -catenin levels in wild type fibroblasts



**Figure 6.2:** Western blot analysis of  $\beta$ -catenin levels in wild type murine embryonic fibroblasts following exposure to MNNG. **A)** 90kDa signal corresponding to  $\beta$ -catenin. Lanes 1-3 are protein extracts from untreated fibroblasts, lanes 4-6 correspond to protein extracts from the same fibroblast preparations treated with 50  $\mu$ M MNNG for 15 hours prior to harvesting. **B)** An example of a coomassie stained acrylamide gel showing the high molecular weight proteins retained in the gel following protein transfer to nitrocellulose membrane; the stained gels were used to correct for small variations in protein loading levels. **C)** Loading corrected densitometry values for the  $\beta$ -catenin signal.

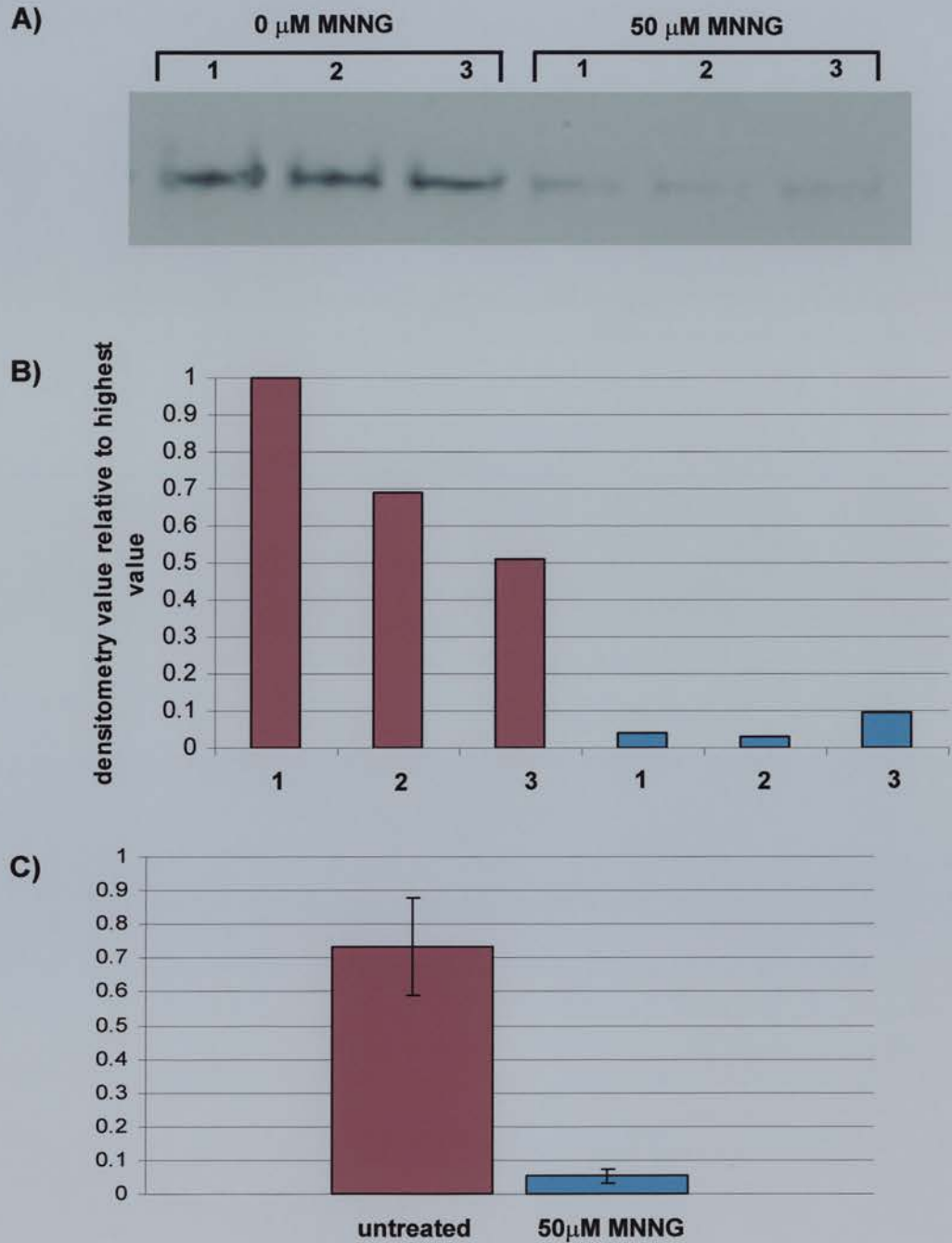
**Figure 6.3:**  $\beta$ -catenin levels in p53 null fibroblasts



**Figure 6.3:** Western blot analysis of  $\beta$ -catenin levels in MSH2 null murine embryonic fibroblast following exposure to MNNG. **A)** 90kDa signal corresponding to  $\beta$ -catenin. Lanes 1-3 are protein extracts from untreated fibroblasts, lanes 4-6 correspond to protein extracts from the same fibroblast preparations treated with 50  $\mu$ M MNNG for 15 hours prior to harvesting. **B)** Loading corrected densitometry values for the  $\beta$ -catenin signal. 1-3 are untreated fibroblast values, 4-6 are MNNG treated. **C)** Averaged corrected values showing a relatively small reduction in the relative  $\beta$ -catenin levels occurring on exposure to MNNG.

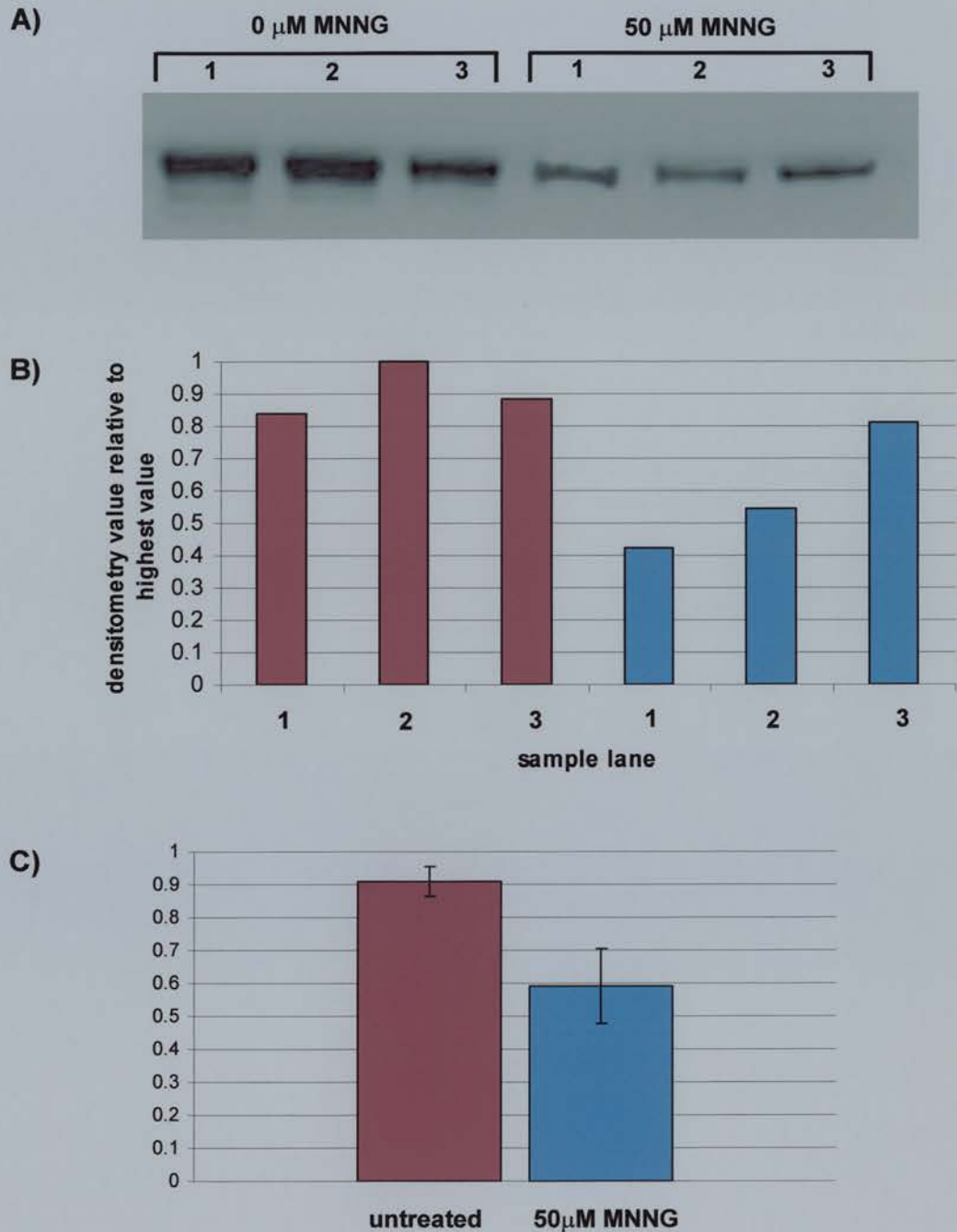


**Figure 6.4:**  $\beta$ -catenin levels in MSH2 null fibroblasts



**Figure 6.4:** Western blot analysis of  $\beta$ -catenin levels in MSH2 null murine embryonic fibroblast following exposure to MNNG. **A)** 90kDa signal corresponding to  $\beta$ -catenin. Lanes 1-3 are protein extracts from untreated fibroblasts, lanes 4-6 correspond to protein extracts from the same fibroblast preparations treated with 50  $\mu$ M MNNG for 15 hours prior to harvesting. **B)** Loading corrected densitometry values for the  $\beta$ -catenin signal. 1-3 are untreated fibroblast values, 4-6 are MNNG treated. **C)** Averaged corrected values showing a large reduction in the relative  $\beta$ -catenin levels occurring on exposure to MNNG.

**Figure 6.5:**  $\beta$ -catenin levels in p53/MSH2 null fibroblasts



**Figure 6.5:** Western blot analysis of  $\beta$ -catenin levels in p53/MSH2 null murine embryonic fibroblast following exposure to MNNG. **A)** 90kDa signal corresponding to  $\beta$ -catenin. Lanes 1-3 are protein extracts from untreated fibroblasts, lanes 4-6 correspond to protein extracts from the same fibroblast preparations treated with 50  $\mu$ M MNNG for 15 hours prior to harvesting. **B)** Loading corrected densitometry values for the  $\beta$ -catenin signal. 1-3 are untreated fibroblast values, 4-6 are MNNG treated. **C)** Averaged corrected values showing a relatively small reduction in the relative  $\beta$ -catenin levels occurring on exposure to MNNG.



## 6.4 Discussion

The upregulation of APC levels in human cell lines and MEFs in response to MNNG exposure has previously been shown to depend on an increase in cellular levels of functional p53 tumour suppressor protein (Jaiswal and Narayan, 1998; Narayan and Jaiswal, 1997). It has also previously been shown that a functional MutS $\alpha$  pathway, which includes the Msh2 protein, is a requirement for the stabilisation and subsequent accumulation of p53 in MEFs in response to *O*<sup>6</sup>MeGs formation following MNNG exposure (Hickman and Samson, 1999). Based on these observations it would then be expected that APC accumulation and the subsequent increased rate of degradation of  $\beta$ -catenin would depend on the presence of functional Msh2.

The data presented here clearly show a reduction in the levels of  $\beta$ -catenin in MNNG treated *Msh2* null fibroblasts, while the levels of this protein remain relatively constant in p53 null or p53/*Msh2* null cells regardless of treatment. It would appear that  $\beta$ -catenin levels are suppressed in a manner at least partially dependent on p53 but independent of *Msh2* status. This suppression may occur either at the level of transcription or post-translationally through some APC-independent breakdown of  $\beta$ -catenin. While some evidence points to a dependence on the presence of a functional MutS $\alpha$  pathway to upregulate and/or phosphorylate p53 in various cell lines, there is strong evidence for some functional independence between *Msh2* and p53 DNA damage responses in murine systems. The administration of 50mg/kg MNNG to wild type mice results in elevated levels of apoptosis in the small intestine. This increase in apoptosis does not appear to occur in p53 null mouse lines, indicating that the apoptotic response to MNNG is p53 dependant. The same chemical treatment of *Msh2* null mouse lines resulted in a reduced but still significant wave of apoptosis some 6 hours following treatment with MNNG (Toft *et al.*, 1999). While p53 stabilisation following MNNG treatment may be strictly dependent on a functional MutS $\alpha$  pathway in some cell lines, there is clearly some evidence for downstream p53 effects following MNNG treatment in MutS $\alpha$  deficient *in vivo* systems.

It is possible that MNNG induced lesions other than  $O^6$ MeGs, or indeed alkylation of molecules other than DNA, are capable of initiating p53 dependant pathways which ultimately cause the observed effects on  $\beta$ -catenin levels. Double strand breaks in DNA are known to induce cellular responses, including apoptosis, in a p53 dependant manner; these breaks may arise in actively dividing cells which have failed to remove DNA adducts prior to DNA replication (Karran and Bignami, 1994). The MEF populations in the set of experiments presented in this chapter were cultured in 0.5% FCS starvation medium prior to and during MNNG treatment in an effort to slow or stop cell division. The *Msh2* null cells, both with and without p53 deficiency, appeared to be able to grow and divide despite the low level of serum supplement in the culture medium. An attempt was made to quantify this phenomenon using clonogenic survival studies (data not presented), however the ability of the *Msh2* null cells to replicate in starvation medium seemed to be highly confluence dependent. When either *Msh2*<sup>-/-</sup> or *Msh2*<sup>-/-</sup>/p53<sup>-/-</sup> null cell at 75% confluence were switched to starvation medium 18 hours prior to MNNG treatment they often reached nearly 100% confluence prior to MNNG addition, while the sparse seeding in the clonogenic survival study produced very little growth, making the determination of number of clones very difficult. The apparent ability of *Msh2* null cells to grow through starvation regimes may allow for the formation of DNA strand breaks and the associated p53 stabilisation. This would in turn lead to *Apc* upregulation and finally  $\beta$ -catenin degradation. The ability of MMR deficient cells to grow through normal cell cycle checkpoints has been observed previously; both *MLH1* null human cell lines and murine fibroblasts have been shown to have a reduced G2/M cell checkpoint arrest following ionising radiation (Davis *et al.*, 1998).

While p53 upregulation of *Apc* levels is perhaps the most likely cause of the observed  $\beta$ -catenin downregulation there are other ways in which the levels of  $\beta$ -catenin present within a cell may be altered without a significant increase in *Apc* protein levels.  $\beta$ -Transducin repeat containing protein ( $\beta$ -Trcp) has recently been shown to interact with phosphorylated  $\beta$ -catenin and can influence its

degradation by the ubiquitin-proteasome pathway (Hart *et al.*, 1999; Liu *et al.*, 1999; Easwaran *et al.*, 1999b). Another protein, casein kinase I $\epsilon$  (CKI $\epsilon$ ), has also been shown to influence  $\beta$ -catenin degradation. This protein binds the Wnt-1 pathway component Axin; overexpression of CKI $\epsilon$  mimics the effects of Wnt signalling, resulting in the accumulation of  $\beta$ -catenin and the activation of transcription from Wnt target genes (Sakanaka *et al.*, 1999). Upregulation of  $\beta$ -Tcrp levels or downregulation of CKI $\epsilon$  could both result in the increased rates of  $\beta$ -catenin degradation observed in the *Msh2* null fibroblasts. It is possible that cellular levels of  $\beta$ -Tcrp or CKI $\epsilon$  may be modulated by p53.

# **Chapter 7**

## **Conclusions**

## Chapter 7 – Conclusions

The investigations undertaken in this thesis focussed on three main areas; the first being the characterisation of novel promoter sequence of the intestinal epithelial-specific *Cdx-1* homeobox gene and work towards engineering this sequence into a conditional transgenic construct, secondly, the thorough phenotypic characterisation of a floxed APC transgenic mouse line and finally the investigation of downstream effects of methylating agent exposure on the Wnt-1 pathway.

Numerous attempts to generate a BAC based *Cdx-1* promoter driven Cre expression transgene through the use of an *in vivo* recombination system failed to produce the desired transgene construct. The failure to generate a recombination modified BAC, despite efforts to optimise the system, could have a number of possible causes, as discussed in section 3.4. The conclusions drawn here are that, while the potential for the rapid generation of transgene or targeting constructs using this *in vivo* system is great, the efficiency of the system may be hampered by a number of factors. Careful consideration of the nature of the selection systems employed, as well as an awareness of potential problems posed by special features of the sequences composing the arms of homology are key to successful application of this *in vivo* recombination system. It may be that certain BAC manipulations will prove inefficient or impossible within the limitations of the Red $\alpha$ - $\beta$ - $\gamma$  based protocol. At least one other BAC alteration system based on recombination has been developed; this system uses a single arm of homology and a highly efficient Cre mediated recombination event to introduce planned changes into the BAC sequence (Dr. Chris Boyd, personal communication). This alternative system may overcome some of the problems posed by the Red $\alpha$ - $\beta$ - $\gamma$  recombination protocol.

A fragment of one of the *Cdx-1* positive BAC clones (BAC 219) containing novel *Cdx-1* promoter sequence was successfully subcloned by Dr. Sula Corbet and Dr. Jane Armstrong. As part of this thesis, this fragment was mapped and its full length sequenced; the derived sequence data revealed a number of putative transcription factor binding sites in addition to confirming the presence of a

number of previously partially characterised motifs. Four EGFP reporter plasmids, driven by varying lengths of the isolated *Cdx-1* promoter sequence, each containing a different complement of transcription factor binding sites, were assembled and introduced into a number of cell types. Expression levels measured varied considerably between the cell types and also between reporter constructs, providing information about transcription motifs responsible for expression levels and possibly tissue specificity.

Had time allowed, an investigation of the role played by the putative transcription factor binding sites in the control of *Cdx-1* transcription would have proved interesting. Preliminary investigations into the extent of transcriptional control of the *Cdx-1* locus contributed by c-myb would consist of DNA footprinting or electrophoretic mobility shift assays (EMSAs) to determine if the sequence motif in question does indeed interact physically with c-myb *in vitro*. Effects of the presence of c-myb on the expression of *Cdx-1* *in vivo* could be conducted by co-transfection of ES cells (or another appropriate cell type) with *Cdx*-EGFP reporters containing the putative c-myb binding site and c-myb overexpression vectors. Alternatively some information regarding the function of the potential c-myb binding site may be obtained through transfection of cell lines known to overexpress c-myb in parallel with the transfection of related lines which do not overexpress this oncogene. Should the results of these experiments suggest a role for c-myb in *Cdx-1* expression, site directed mutagenesis of the c-myb binding site could be performed to determine if sequence alterations abrogate the previously observed c-myb effects. Similar experiments could be performed to determine if the putative CREB binding sites identified are physiologically important transcription factor motifs. The investigation of the properties of the isolated *Cdx-1* promoter sequence using *Cdx-1* based reporter constructs would be of interest in the study of developmental control of the *Cdx-1* homeobox gene and will be of importance if characterisation of the *Cdx-1*-CreER<sup>tm</sup> transgenic mouse line generated by Drs. Jane Armstrong and Sula Corbet reveals any unexpected pattern of expression or recombination.

Phenotypic characterisation of the *CC* floxed APC mouse line revealed that, while some expression from the introduced transgenic construct could be



detected by the very sensitive RT-PCR method, presence of the transgene was insufficient to significantly alter the *Min* phenotype when the transgenic line was crossed onto the *Apc* mutant strain. Histological appearances, embryogenesis, gross anatomy, survival times and tumour burden were not significantly altered by the presence of the pAPC 2M35 transgene. The lack of phenotype alteration seen in the *CC* line could have a number of causes. Perhaps the most plausible explanation for the lack of effect is absence of sufficient expression. Low levels of expression from the pAPC 2M35 transgene could have a number of causes, including poor expression from the PGK promoter used in the transgene construct or the integration of a small number of copies of the transgene into a silenced area of the mouse genome. While it was possible to establish through Southern hybridisation that the transgene was in fact present in the genome in a detectable quantity, no information regarding the number of integration sites has been gleaned from this hybridisation. Additionally, Southern detection was only accomplished using genomic DNA from later generations of *CC* mice which may have inherited copies of the transgene array from both transgene positive parents, effectively doubling the transgene copy number. Inheritance of the transgene in early breeding from the founder mouse, and the generation immediately following, suggested that the transgene had integrated in only one location within the single founder animal's genome. The difficulty in detecting the transgene in early mice, coupled with the breeding information regarding transgene inheritance, suggests that the transgene was present in a relatively small number of copies in only one genomic location. This increases the likelihood of low expression levels, caused by integration into a silenced location, being the cause for the lack of rescue of the *Min* phenotype.

While the transgene has thus far failed to modify the *Min* phenotype, the transfection of the transgene into human colorectal derived cell lines appears to show that the transgene is transcribed *in vitro*. Further pronuclear injections may generate a founder mouse with multiple sites of transgene array integration, possibly generating sufficient expression to affect the *Min* phenotype. Additionally, it may be possible to re-engineer the pAPC 2M35 transgene construct to increase transcription levels; a different promoter may give higher or more consistent expression across different cell types and replacing the SV40

polyA signal with the bovine growth hormone polyA signal may result in significantly higher translation levels (Pfarr *et al.*, 1986).

The final area of investigation presented in this thesis involves the Wnt-1 pathway responses to the alkylating agent MNNG in MEFs derived from mice engineered to be deficient for the tumour suppressor genes p53 and/or *Msh2*. The upregulation of *Apc* in response to exposure to MNNG has been observed in a human derived colorectal cell line and in MEFs and is dependent on the presence of functional p53 protein. There is evidence to suggest that the stabilisation of p53 in human lymphoblastoma cell lines exposed to MNNG is dependent on a functional MutS $\alpha$  pathway and as such is *Msh2* dependent. The data presented here shows a clear downregulation of the *Apc* target  $\beta$ -catenin in MEFs deficient in *Msh2* following 15 hours of exposure to 50 $\mu$ m MNNG, while MEFs derived from p53 null animals showed a much smaller reduction in  $\beta$ -catenin following the same treatment. This suggests that the downregulation of  $\beta$ -catenin in response to MNNG is p53 dependent but at least partially independent of the MutS $\alpha$  pathway.

Evidence for *Msh2* independent p53 effects following MNNG treatment has been seen previously *in vivo*. In wildtype mice treated with 50mg/kg MNNG there is a wave of p53 dependent apoptosis seen in the small intestine occurring 6 hours following treatment. In *Msh2* deficient mice this response is reduced by approximately 50-60% but still significant, indicating that there exists some MutS $\alpha$  independent means of activating p53 following MNNG treatment (Toft *et al.*, 1999). The p53 dependent / *Msh2* independent reduction in  $\beta$ -catenin levels in MNNG treated fibroblasts may be triggered in a manner similar to that occurring in the mouse small intestine.

Time permitting, further investigations of the observed reduction in  $\beta$ -catenin levels following MNNG treatment would have been carried out. It would be of interest to determine if p53 accumulation could be detected by immunoblot in the *Msh2* null MNNG treated MEFs. Failure to detect increased amounts of p53 by western blot may indicate that the p53 dependent MNNG response displayed by

the MEFs is transduced through some means other than a simple increase in the amount of the p53 protein present in the cell, while increased levels of p53 would point to an *Msh2* independent stabilisation of p53 in these cells. Also of interest would be an investigation of the kinetics of  $\beta$ -catenin loss following MNNG exposure, coupled with a measure of cell proliferation or cycling. A time course experiment may show a delay in the  $\beta$ -catenin reduction in the *Msh2* null cells when compared with the wild type MEFs which, if coupled with a detection of persistent proliferation of *Msh2* null fibroblasts in starvation medium, would begin to suggest that failure to remove methyl adducts prior to DNA synthesis is involved in the observed downstream p53 effects.

The upregulation of APC levels in response to MNNG was initially observed in a human colorectal cell line containing wild type *APC* alleles but a degradation resistant  $\beta$ -catenin mutation. The measurement of  $\beta$ -catenin levels in APC deficient cell lines (such as SW 480) following MNNG exposure would help to determine if the observed reduction of  $\beta$ -catenin *in vitro* on MNNG exposure is solely (or largely) dependent on an increase in APC.

The development of improved and refined conditional transgenic and gene targeting techniques will undoubtedly contribute much to our understanding of biological processes, including the early stages of colorectal carcinogenesis. Ligand activation of inducible promoters or fusion proteins will continue to improve, offering tighter control over the timing of gene expression. At the same time, the number of known promoter sequences, including novel tissue-specific promoters, will only grow as more genomes are sequenced and analysed. Both of these developments will enable the dissection of molecular events in normal development and of genetic dysregulation in pathogenesis.

# References

- Aaltonen, L.A. and Peltomaki, P.(1994). Genes involved in hereditary nonpolyposis colorectal carcinoma. *Anticancer Res* **14**, 1657-1660.
- Aaltonen, L.A., Peltomaki, P., Leach, F.S., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Powell, S.M., Jen, J., and Hamilton, S.R.(1993). Clues to the pathogenesis of familial colorectal cancer [see comments]. *Science* **260**, 812-816.
- Aberle, H., Bauer, A., Stappert, J., Kispert, A., and Kemler, R.(1997). beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J* **16**, 3797-3804.
- Aberle, H., Schwartz, H., and Kemler, R.(1996). Cadherin-catenin complex: protein interactions and their implications for cadherin function. *J Cell Biochem* **61**, 514-523.
- Abremski, K. and Hoess, R.(1984). Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J Biol Chem* **259**, 1509-1514.
- Abremski, K., Hoess, R., and Sternberg, N.(1983). Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell* **32**, 1301-1311.
- Abuin, A. and Bradley, A.(1996). Recycling selectable markers in mouse embryonic stem cells. *Mol Cell Biol* **16**, 1851-1856.
- Adra, C.N., Boer, P.H., and McBurney, M.W.(1987). Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene* **60**, 65-74.
- Agah, R., Frenkel, P.A., French, B.A., Michael, L.H., Overbeek, P.A., and Schneider, M.D.(1997). Gene recombination in postmitotic cells. Targeted expression of Cre recombinase provokes cardiac-restricted, site-specific rearrangement in adult ventricular muscle in vivo. *J Clin Invest* **100**, 169-179.
- Alitalo, K., Winqvist, R., Lin, C.C., de la Chapelle, A., Schwab, M., and Bishop, J.M.(1984). Aberrant expression of an amplified c-myc oncogene in two cell lines from a colon carcinoma. *Proc Natl Acad Sci U S A* **81**, 4534-4538.
- Amin, A., Roca, H., Luetke, K., and Sadowski, P.D.(1991). Synapsis, strand scission, and strand exchange induced by the FLP recombinase: analysis with half-FRT sites. *Mol Cell Biol* **11**, 4497-4508.
- Amin, A.A., Beatty, L.G., and Sadowski, P.D.(1990). Synaptic intermediates promoted by the FLP recombinase. *J Mol Biol* **214**, 55-72.
- Andrews, B.J., Beatty, L.G., and Sadowski, P.D.(1987). Isolation of intermediates in the binding of the FLP recombinase of the yeast plasmid 2-micron circle to its target sequence. *J Mol Biol* **193**, 345-358.
- Anton, M. and Graham, F.L.(1995). Site-specific recombination mediated by an adenovirus vector expressing the Cre recombinase protein: a molecular switch for control of gene expression. *J Virol* **69**, 4600-4606.
- Aoki, M., Hecht, A., Kruse, U., Kemler, R., and Vogt, P.K.(1999). Nuclear endpoint of Wnt signaling: neoplastic transformation induced by transactivating lymphoid-enhancing factor 1. *Proc Natl Acad Sci U S A* **96**, 139-144.
- Araki, K., Araki, M., and Yamamura, K.(1997). Targeted integration of DNA using mutant lox sites in embryonic stem cells. *Nucleic Acids Res* **25**, 868-872.
- Askew, G.R., Doetschman, T., and Lingrel, J.B.(1993). Site-directed point mutations in embryonic stem cells: a gene-targeting tag-and-exchange strategy. *Mol Cell Biol* **13**, 4115-4124.

- Axelrod, J.D., Miller, J.R., Shulman, J.M., Moon, R.T., and Perrimon, N.(1998). Differential recruitment of Dishevelled provides signaling specificity in the planar cell polarity and Wntless signaling pathways. *Genes Dev* **12**, 2610-2622.
- Bach, S.P., Renehan, A.G., and Potten, C.S.(2000). Stem cells: the intestinal stem cell as a paradigm. *Carcinogenesis* **21**, 469-476.
- Bardos, J., Sulekova, Z., and Ballhausen, W.G.(1997). Novel exon connections of the brain-specific (BS) exon of the adenomatous polyposis coli gene. *Int J Cancer* **73**, 137-142.
- Barker, N., Morin, P.J., and Clevers, H.(2000). The Yin-Yang of TCF/beta-catenin signaling. *Adv Cancer Res* **77**, 1-24.
- Barth, A.I., Pollack, A.L., Altschuler, Y., Mostov, K.E., and Nelson, W.J.(1997). NH2-terminal deletion of beta-catenin results in stable colocalization of mutant beta-catenin with adenomatous polyposis coli protein and altered MDCK cell adhesion. *J Cell Biol* **136**, 693-706.
- Behrens, J.(1999). Cadherins and catenins: role in signal transduction and tumor progression. *Cancer Metastasis Rev* **18**, 15-30.
- Behrens, J., Jerchow, B.A., Würtele, M., Grimm, J., Asbrand, C., Wirtz, R., Kühl, M., Wedlich, D., and Birchmeier, W.(1998). Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science* **280**, 596-599.
- Behrens, J., von Kries, J.P., Kuhl, M., Bruhn, L., Wedlich, D., Grosschedl, R., and Birchmeier, W.(1996). Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature* **382**, 638-642.
- Beinhauer, J.D., Hagan, I.M., Hegemann, J.H., and Fleig, U.(1997). Mal3, the fission yeast homologue of the human APC-interacting protein EB-1 is required for microtubule integrity and the maintenance of cell form. *J Cell Biol* **139**, 717-728.
- Beroud, C. and Soussi, T.(1996). APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Res* **24**, 121-124.
- Bethke, B. and Sauer, B.(1997). Segmental genomic replacement by Cre-mediated recombination: genotoxic stress activation of the p53 promoter in single-copy transformants. *Nucleic Acids Res* **25**, 2828-2834.
- Betz, U.A., Voshchenrich, C.A., Rajewsky, K., and Muller, W.(1996). Bypass of lethality with mosaic mice generated by Cre-loxP-mediated recombination. *Curr Biol* **6**, 1307-1316.
- Bhanot, P., Brink, M., Samos, C.H., Hsieh, J.C., Wang, Y., Macke, J.P., Andrew, D., Nathans, J., and Nusse, R.(1996). A new member of the frizzled family from Drosophila functions as a Wntless receptor. *Nature* **382**, 225-230.
- Bhat, R.V., Baraban, J.M., Johnson, R.C., Eipper, B.A., and Mains, R.E.(1994). High levels of expression of the tumor suppressor gene APC during development of the rat central nervous system. *J Neurosci* **14**, 3059-3071.
- Bhattacharjee, R.N., Hamada, F., Toyoshima, K., and Akiyama, T.(1996). The tumor suppressor gene product APC is hyperphosphorylated during the M phase. *Biochem Biophys Res Commun* **220**, 192-195.
- Bilger, A., Shoemaker, A.R., Gould, K.A., and Dove, W.F.(1996). Manipulation of the mouse germline in the study of Min-induced neoplasia. *Semin Cancer Biol* **7**, 249-260.



- Boland, C.R., Sato, J., Saito, K., Carethers, J.M., Marra, G., Laghi, L., and Chauhan, D.P.(1998). Genetic instability and chromosomal aberrations in colorectal cancer: a review of the current models. *Cancer Detect Prev* **22**, 377-382.
- Bollag, R.J., Waldman, A.S., and Liskay, R.M.(1989). Homologous recombination in mammalian cells. *Annu Rev Genet* **23**, 199-225.
- Booth, C., Patel, S., Bennion, G.R., and Potten, C.S.(1995). The isolation and culture of adult mouse colonic epithelium. *Epithelial Cell Biol* **4**, 76-86.
- Bouhassira, E.E., Westerman, K., and Leboulch, P.(1997). Transcriptional behavior of LCR enhancer elements integrated at the same chromosomal locus by recombinase-mediated cassette exchange. *Blood* **90**, 3332-3344.
- Brabletz, T., Jung, A., Dag, S., Hlubek, F., and Kirchner, T.(1999). beta-catenin regulates the expression of the matrix metalloproteinase-7 in human colorectal cancer. *Am J Pathol* **155**, 1033-1038.
- Brattain, M.G., Fine, W.D., Khaled, F.M., Thompson, J., and Brattain, D.E.(1981). Heterogeneity of malignant cells from a human colonic carcinoma. *Cancer Res* **41**, 1751-1756.
- Broach, J.R. and Hicks, J.B.(1980). Replication and recombination functions associated with the yeast plasmid, 2 mu circle. *Cell* **21**, 501-508.
- Brocard, J., Warot, X., Wendling, O., Messaddeq, N., Vonesch, J.L., Chambon, P., and Metzger, D.(1997). Spatio-temporally controlled site-specific somatic mutagenesis in the mouse. *Proc Natl Acad Sci U S A* **94**, 14559-14563.
- Brooks, A.I., Muhkerjee, B., Panahian, N., Cory-Slechta, D., and Federoff, H.J.(1997). Nerve growth factor somatic mosaicism produced by herpes virus-directed expression of cre recombinase. *Nat Biotechnol* **15**, 57-62.
- Browne, S.J., MacFarlane, M., Cohen, G.M., and Paraskeva, C.(1998). The adenomatous polyposis coli protein and retinoblastoma protein are cleaved early in apoptosis and are potential substrates for caspases. *Cell Death Differ* **5**, 206-213.
- Browne, S.J., Williams, A.C., Hague, A., Butt, A.J., and Paraskeva, C.(1994). Loss of APC protein expressed by human colonic epithelial cells and the appearance of a specific low-molecular-weight form is associated with apoptosis in vitro. *Int J Cancer* **59**, 56-64.
- Burnstein, M.J.(1993). Dietary factors related to colorectal neoplasms. *Surg Clin North Am* **73**, 13-29.
- Burt, R.W., Bishop, D.T., Cannon-Albright, L., Samowitz, W.S., Lee, R.L., DiSario, J.A., and Skolnick, M.H.(1992). Population genetics of colonic cancer. *Cancer* **70**, 1719-1722.
- Butz, S. and Kemler, R.(1994). Distinct cadherin-catenin complexes in Ca(2+)-dependent cell-cell adhesion. *FEBS Lett* **355**, 195-200.
- Butz, S. and Larue, L.(1995). Expression of catenins during mouse embryonic development and in adult tissues. *Cell Adhes Commun* **3**, 337-352.
- Cabrera, C.V., Alonso, M.C., Johnston, P., Phillips, R.G., and Lawrence, P.A.(1987). Phenocopies induced with antisense RNA identify the wingless gene. *Cell* **50**, 659-663.
- Cadigan, K.M. and Nusse, R.(1997). Wnt meeting 1996. *Biochim Biophys Acta* **1332**, R1-R5

- Calvert, R. and Pothier, P.(1990). Migration of fetal intestinal intervillous cells in neonatal mice. *Anat Rec* **227**, 199-206.
- Cannon-Albright, L.A., Bishop, D.T., Goldgar, C., and Skolnick, M.H.(1991). Genetic predisposition to cancer. *Important Adv Oncol* 39-55.
- Cannon-Albright, L.A., Skolnick, M.H., Bishop, D.T., Lee, R.G., and Burt, R.W.(1988). Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. *N Engl J Med* **319**, 533-537.
- Carter, C.A., Pogribny, M., Davidson, A., Jackson, C.D., McGarrity, L.J., and Morris, S.M.(1996). Effects of retinoic acid on cell differentiation and reversion toward normal in human endometrial adenocarcinoma (RL95-2) cells. *Anticancer Res* **16**, 17-24.
- Casciola-Rosen, L., Nicholson, D.W., Chong, T., Rowan, K.R., Thornberry, N.A., Miller, D.K., and Rosen, A.(1996). Apopain/CPP32 cleaves proteins that are essential for cellular repair: a fundamental principle of apoptotic death [see comments]. *J Exp Med* **183**, 1957-1964.
- Chouluka, A., Guyot, V., and Nicolas, J.F.(1996). Transfer of single gene-containing long terminal repeats into the genome of mammalian cells by a retroviral vector carrying the cre gene and the loxP site. *J Virol* **70**, 1792-1798.
- Chow, J.S., Chen, C.C., Ahsan, H., and Neugut, A.I.(1996). A population-based study of the incidence of malignant small bowel tumours: SEER, 1973-1990. *Int J Epidemiol* **25**, 722-728.
- Cillo, C., Barba, P., Freschi, G., Bucciarelli, G., Magli, M.C., and Boncinelli, E.(1992). HOX gene expression in normal and neoplastic human kidney. *Int J Cancer* **51**, 892-897.
- Clarke, A.R., Gledhill, S., Hooper, M.L., Bird, C.C., and Wyllie, A.H.(1994). p53 dependence of early apoptotic and proliferative responses within the mouse intestinal epithelium following gamma-irradiation. *Oncogene* **9**, 1767-1773.
- Clarke, A.R., Maandag, E.R., van Roon, M., van der Lugt, N.M., van der Valk, M., Hooper, M.L., Berns, A., and te, R.(1992). Requirement for a functional Rb-1 gene in murine development [see comments]. *Nature* **359**, 328-330.
- Clarke, A.R., Purdie, C.A., Harrison, D.J., Morris, R.G., Bird, C.C., Hooper, M.L., and Wyllie, A.H.(1993). Thymocyte apoptosis induced by p53-dependent and independent pathways [see comments]. *Nature* **362**, 849-852.
- Clevers, H. and van de Wetering, M.(1997). TCF/LEF factor earn their wings. *Trends Genet* **13**, 485-489.
- Cooper, C.(2000). Clinical and experimental studies on the role of APC in neoplasia. *PhD Thesis* (in press)
- Cormack, B.P., Valdivia, R.H., and Falkow, S.(1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-38.
- Cormier, R.T., Hong, K.H., Halberg, R.B., Hawkins, T.L., Richardson, P., Mulharker, R., Dove, W.F., and Lander, E.S.(1997). Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis [see comments]. *Nat Genet* **17**, 88-91.
- Crawford, H.C., Fingleton, B.M., Rudolph-Owen, L.A., Goss, K.J., Rubinfeld, B., Polakis, P., and Matrisian, L.M.(1999). The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors. *Oncogene* **18**, 2883-2891.

- Danielian, P.S., Muccino, D., Rowitch, D.H., Michael, S.K., and McMahon, A.P.(1998). Modification of gene activity in mouse embryos in utero by a tamoxifen-inducible form of Cre recombinase. *Curr Biol* **8**, 1323-1326.
- Davidson, H., Doherty, A., Stevenson, B., Boyd, A.C., and Porteous, D.J.(1999). Elimination of contaminant Escherichia coli chromosomal DNA from preparations of P1 artificial chromosome recombinants facilitates directed subcloning. *Electrophoresis* **20**, 1469-1475.
- Davies, D.R., Armstrong, J.G., Thakker, N., Horner, K., Guy, S.P., Clancy, T., Sloan, P., Blair, V., Dodd, C., and Warnes, T.W.(1995). Severe Gardner syndrome in families with mutations restricted to a specific region of the APC gene. *Am J Hum Genet* **57**, 1151-1158.
- Davis, T.W., Wilson-Van Patten, C., Meyers, M., Kunugi, K.A., Cuthill, S., Reznikoff, C., Garces, C., Boland, C.R., Kinsella, T.J., Fishel, R., and Boothman, D.A.(1998). Defective expression of the DNA mismatch repair protein, MLH1, alters G2-M cell cycle checkpoint arrest following ionizing radiation. *Cancer Res* **58**, 767-778.
- de La Coste, Romagnolo, B., Billuart, P., Renard, C.A., Buendia, M.A., Soubrane, O., Fabre, M., Chelly, J., Beldjord, C., Kahn, A., and Perret, C.(1998). Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proc Natl Acad Sci U S A* **95**, 8847-8851.
- Deka, J., Herter, P., Sprenger-Haussels, M., Koosch, S., Franz, D., Muller, K.M., Kuhnen, C., Hoffmann, I., and Muller, O.(1999). The APC protein binds to A/T rich DNA sequences. *Oncogene* **18**, 5654-5661.
- Del Bufalo, D., Cucco, C., Leonetti, C., Citro, G., D'Agnano, I., Benassi, M., Geiser, T., Zon, G., Calabretta, B., and Zupi, G.(1996). Effect of cisplatin and c-myc antisense phosphorothioate oligodeoxynucleotides combination on a human colon carcinoma cell line in vitro and in vivo. *Br J Cancer* **74**, 387-393.
- Dexter, D.L. and Hager, J.C.(1980). Maturation-induction of tumor cells using a human colon carcinoma model. *Cancer* **45**, 1178-1184.
- Dias, J.M., Go, N.F., Hart, C.P., and Mattheakis, L.C.(1998). Genetic recombination as a reporter for screening steroid receptor agonists and antagonists. *Anal Biochem* **258**, 96-102.
- Dietrich, W.F., Lander, E.S., Smith, J.S., Moser, A.R., Gould, K.A., Luongo, C., Borenstein, N., and Dove, W.(1993). Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. *Cell* **75**, 631-639.
- Dihlmann, S., Amler, L.C., Schwab, M., and Wenzel, A.(1997). Variations in the expression of the adenomatous polyposis coli (APC) tumor suppressor gene in human cancer cell lines of different tissue origin. *Oncol Res* **9**, 119-127.
- Dihlmann, S., Gebert, J., Siermann, A., Herfarth, C., and von Knebel, D.(1999). Dominant negative effect of the APC1309 mutation: a possible explanation for genotype-phenotype correlations in familial adenomatous polyposis. *Cancer Res* **59**, 1857-1860.
- Dobashi, Y., Bhattacharjee, R.N., Toyoshima, K., and Akiyama, T.(1996). Upregulation of the APC gene product during neuronal differentiation of rat pheochromocytoma PC12 cells. *Biochem Biophys Res Commun* **224**, 479-483.
- Dominguez, I., Itoh, K., and Sokol, S.Y.(1995). Role of glycogen synthase kinase 3 beta as a negative regulator of dorsoventral axis formation in Xenopus embryos. *Proc Natl Acad Sci U S A* **92**, 8498-8502.

- Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A.J., Butel, J.S., and Bradley, A.(1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215-221.
- Dorin, J.R.(1995). Development of mouse models for cystic fibrosis. *J Inherit Metab Dis* **18**, 495-500.
- Dorin, J.R., Dickinson, P., Emslie, E., Clarke, A.R., Dobbie, L., Hooper, M.L., Halford, S., Wainwright, B.J., and Porteous, D.J.(1992). Successful targeting of the mouse cystic fibrosis transmembrane conductance regulator gene in embryonal stem cells. *Transgenic Res* **1**, 101-105.
- Dou, Q.P., An, B., and Will, P.L.(1995). Induction of a retinoblastoma phosphatase activity by anticancer drugs accompanies p53-independent G1 arrest and apoptosis. *Proc Natl Acad Sci U S A* **92**, 9019-9023.
- Duluc, I., Lorentz, O., Fritsch, C., Leberquier, C., Kedinger, M., and Freund, J.N.(1997). Changing intestinal connective tissue interactions alters homeobox gene expression in epithelial cells. *J Cell Sci* **110** ( Pt 11), 1317-1324.
- Dyson, P.J., Poirier, F., and Watson, R.J.(1989). Expression of c-myc in embryonal carcinoma cells and embryonal stem cells. *Differentiation* **42**, 24-27.
- Easwaran, V., Pishvaian, M., Salimuddin, and Byers, S.(1999a). Cross-regulation of beta-catenin-LEF/TCF and retinoid signaling pathways. *Curr Biol* **9**, 1415-1418.
- Easwaran, V., Song, V., Polakis, P., and Byers, S.(1999b). The ubiquitin-proteasome pathway and serine kinase activity modulate adenomatous polyposis coli protein-mediated regulation of beta-catenin-lymphocyte enhancer-binding factor signaling. *J Biol Chem* **274**, 16641-16645.
- Evans, M.J. and Kaufman, M.H.(1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
- Fagotto, F., Gluck, U., and Gumbiner, B.M.(1998). Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. *Curr Biol* **8**, 181-190.
- Fagotto, F., Guger, K., and Gumbiner, B.M.(1997). Induction of the primary dorsalizing center in *Xenopus* by the Wnt/GSK/beta-catenin signaling pathway, but not by Vg1, Activin or Noggin. *Development* **124**, 453-460.
- Fagotto, F., Jho, E., Zeng, L., Kurth, T., Joos, T., Kaufmann, C., and Costantini, F.(1999). Domains of axin involved in protein-protein interactions, Wnt pathway inhibition, and intracellular localization. *J Cell Biol* **145**, 741-756.
- Feil, R., Brocard, J., Mascrez, B., LeMeur, M., Metzger, D., and Chambon, P.(1996). Ligand-activated site-specific recombination in mice. *Proc Natl Acad Sci U S A* **93**, 10887-10890.
- Feil, R., Wagner, J., Metzger, D., and Chambon, P.(1997). Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem Biophys Res Commun* **237**, 752-757.
- Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., and Kolodner, R.(1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer [published erratum appears in *Cell* 1994 Apr 8;77(1):167]. *Cell* **75**, 1027-1038.
- Fogh, J.Human tumor cells in vitro. 115-159 (1975).(Generic)

- Fukuchi, T., Sakamoto, M., Tsuda, H., Maruyama, K., Nozawa, S., and Hirohashi, S.(1998). Beta-catenin mutation in carcinoma of the uterine endometrium. *Cancer Res* **58**, 3526-3528.
- Fukushige, S. and Sauer, B.(1992). Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *Proc Natl Acad Sci U S A* **89**, 7905-7909.
- Furuuchi, K., Tada, M., Yamada, H., Kataoka, A., Furuuchi, N., Hamada, J., Takahashi, M., Todo, S., and Moriuchi, T.(2000). Somatic mutations of the APC gene in primary breast cancers. *Am J Pathol* **156**, 1997-2005.
- Gamallo, C., Palacios, J., Moreno, G., Calvo, d.M., Suarez, A., and Armas, A.(1999). beta-catenin expression pattern in stage I and II ovarian carcinomas : relationship with beta-catenin gene mutations, clinicopathological features, and clinical outcome. *Am J Pathol* **155**, 527-536.
- Gilbert, S.F.(1994). *Developmental Biology*, Fourth edition. Sinauer Associates Inc., Sunderland, Massachusetts.
- Giovannucci, E. and Willett, W.C.(1994). Dietary factors and risk of colon cancer. *Ann Med* **26**, 443-452.
- Goldstein, M.A. and Doi, R.H.(1995). Prokaryotic promoters in biotechnology. *Biotechnol Annu Rev* **1**, 105-128.
- Gossen, M. and Bujard, H.(1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **89**, 5547-5551.
- Gossler, A., Doetschman, T., Korn, R., Serfling, E., and Kemler, R.(1986). Transgenesis by means of blastocyst-derived embryonic stem cell lines. *Proc Natl Acad Sci U S A* **83**, 9065-9069.
- Gould, K.A., Dietrich, W.F., Borenstein, N., Lander, E.S., and Dove, W.F.(1996a). Mom1 is a semi-dominant modifier of intestinal adenoma size and multiplicity in Min/+ mice. *Genetics* **144**, 1769-1776.
- Gould, K.A., Luongo, C., Moser, A.R., McNeley, M.K., Borenstein, N., Shedlovsky, A., Dove, W.F., Hong, K., Dietrich, W.F., and Lander, E.S.(1996b). Genetic evaluation of candidate genes for the Mom1 modifier of intestinal neoplasia in mice. *Genetics* **144**, 1777-1785.
- Gradl, D., Kuhl, M., and Wedlich, D.(1999). The Wnt/Wg signal transducer beta-catenin controls fibronectin expression. *Mol Cell Biol* **19**, 5576-5587.
- Groden, J., Gelbert, L., Thliveris, A., Nelson, L., Robertson, M., Joslyn, G., Samowitz, W., Spirio, L., Carlson, M., and Burt, R.(1993). Mutational analysis of patients with adenomatous polyposis: identical inactivating mutations in unrelated individuals. *Am J Hum Genet* **52**, 263-272.
- Groden, J., Joslyn, G., Samowitz, W., Jones, D., Bhattacharyya, N., Spirio, L., Thliveris, A., Robertson, M., Egan, S., and Meuth, M.(1995). Response of colon cancer cell lines to the introduction of APC, a colon-specific tumor suppressor gene. *Cancer Res* **55**, 1531-1539.
- Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., and Robertson, M.(1991). Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **66**, 589-600.



- Gu, H., Marth, J.D., Orban, P.C., Mossmann, H., and Rajewsky, K.(1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting [see comments]. *Science* **265**, 103-106.
- Gumbiner, B.M. and McCrea, P.D.(1993). Catenins as mediators of the cytoplasmic functions of cadherins. *J Cell Sci Suppl* **17**, 155-158.
- Guo, F., Gopaul, D.N., and Van Duyne, G.D.(1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* **389**, 40-46.
- Hakem, R., de la Pompa, J.L., Sirard, C., Mo, R., Woo, M., Hakem, A., Wakeham, A., Potter, J., Reitmair, A., Billia, F., Firpo, E., Hui, C.C., Roberts, J., Rossant, J., and Mak, T.W.(1996). The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. *Cell* **85**, 1009-1023.
- Hamada, F., Tomoyasu, Y., Takatsu, Y., Nakamura, M., Nagai, S., Suzuki, A., Fujita, F., Shibuya, H., Toyoshima, K., Ueno, N., and Akiyama, T.(1999). Negative regulation of Wingless signaling by D-axin, a Drosophila homolog of axin. *Science* **283**, 1739-1742.
- Hamilton, D.L. and Abremski, K.(1984). Site-specific recombination by the bacteriophage P1 lox-Cre system. Cre-mediated synapsis of two lox sites. *J Mol Biol* **178**, 481-486.
- Han, M.(1997). Gut reaction to Wnt signaling in worms [comment]. *Cell* **90**, 581-584.
- Hansen, L.A., Sigman, C.C., Andreola, F., Ross, S.A., Kelloff, G.J., and De Luca, L.M.(2000). Retinoids in chemoprevention and differentiation therapy. *Carcinogenesis* **21**, 1271-1279.
- Hart, M., Concordet, J.P., Lassot, I., Albert, I., del los, S., Durand, H., Perret, C., Rubinfeld, B., Margottin, F., Benarous, R., and Polakis, P.(1999). The F-box protein beta-TrCP associates with phosphorylated beta-catenin and regulates its activity in the cell. *Curr Biol* **9**, 207-210.
- Hart, M.J., de los, S., Albert, I.N., Rubinfeld, B., and Polakis, P.(1998). Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. *Curr Biol* **8**, 573-581.
- Hatzfeld, M.(1999). The armadillo family of structural proteins. *Int Rev Cytol* **186**, 179-224.
- Hayashi, S., Rubinfeld, B., Souza, B., Polakis, P., Wieschaus, E., and Levine, A.J.(1997). A Drosophila homolog of the tumor suppressor gene adenomatous polyposis coli down-regulates beta-catenin but its zygotic expression is not essential for the regulation of Armadillo. *Proc Natl Acad Sci U S A* **94**, 242-247.
- He, T.C., Sparks, A.B., Rago, C., Hermeking, H., Zawel, L., da Costa, L.T., Morin, P.J., Vogelstein, B., and Kinzler, K.W.(1998). Identification of c-MYC as a target of the APC pathway [see comments]. *Science* **281**, 1509-1512.
- He, X., Saint-Jeannet, J.P., Woodgett, J.R., Varmus, H.E., and Dawid, I.B.(1995). Glycogen synthase kinase-3 and dorsoventral patterning in Xenopus embryos [published erratum appears in Nature 1995 May 18;375(6528):253]. *Nature* **374**, 617-622.
- Heslip, T.R., Theisen, H., Walker, H., and Marsh, J.L.(1997). Shaggy and dishevelled exert opposite effects on Wingless and Decapentaplegic expression and on positional identity in imaginal discs. *Development* **124**, 1069-1078.
- Hickman, M.J. and Samson, L.D.(1999). Role of DNA mismatch repair and p53 in signaling induction of apoptosis by alkylating agents. *Proc Natl Acad Sci U S A* **96**, 10764-10769.



- Hirschl, D., Bayer, P., and Muller, O.(1996). Secondary structure of an armadillo single repeat from the APC protein. *FEBS Lett* **383**, 31-36.
- Hoess, R., Wierzbicki, A., and Abremski, K.(1987). Isolation and characterization of intermediates in site-specific recombination. *Proc Natl Acad Sci U S A* **84**, 6840-6844.
- Hoess, R.H. and Abremski, K.(1985). Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J Mol Biol* **181**, 351-362.
- Hoess, R.H., Ziese, M., and Sternberg, N.(1982). P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc Natl Acad Sci U S A* **79**, 3398-3402.
- Holliday, R.(1964). A mechanism for gene conversion in fungi. *Genet Res* **5**, 282-304.
- Horii, A., Nakatsuru, S., Ichii, S., Nagase, H., and Nakamura, Y.(1993). Multiple forms of the APC gene transcripts and their tissue-specific expression. *Hum Mol Genet* **2**, 283-287.
- Houle, M., Prinos, P., Iulianella, A., Bouchard, N., and Lohnes, D.(2000). Retinoic acid regulation of Cdx1: an indirect mechanism for retinoids and vertebral specification. *Mol Cell Biol* **20**, 6579-6586.
- Houlston, R.S., Collins, A., Slack, J., and Morton, N.E.(1992a). Dominant genes for colorectal cancer are not rare. *Ann Hum Genet* **56** ( Pt 2), 99-103.
- Houlston, R.S., Fallon, T., Harocopos, C., Williams, C.B., Davey, C., and Slack, J.(1992b). Congenital hypertrophy of retinal pigment epithelium in patients with colonic polyps associated with cancer family syndrome. *Clin Genet* **42**, 16-18.
- Howe, L.R., Subbaramaiah, K., Chung, W.J., Dannenberg, A.J., and Brown, A.M.(1999). Transcriptional activation of cyclooxygenase-2 in Wnt-1-transformed mouse mammary epithelial cells. *Cancer Res* **59**, 1572-1577.
- Hsu, W., Zeng, L., and Costantini, F.(1999). Identification of a domain of Axin that binds to the serine/threonine protein phosphatase 2A and a self-binding domain. *J Biol Chem* **274**, 3439-3445.
- Hu, Y., Kazenwadel, J., and James, R.(1993). Isolation and characterization of the murine homeobox gene Cdx-1. Regulation of expression in intestinal epithelial cells. *J Biol Chem* **268**, 27214-27225.
- Huber, A.H., Nelson, W.J., and Weis, W.I.(1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**, 871-882.
- Huber, O., Bierkamp, C., and Kemler, R.(1996a). Cadherins and catenins in development. *Curr Opin Cell Biol* **8**, 685-691.
- Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B.G., and Kemler, R.(1996b). Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. *Mech Dev* **59**, 3-10.
- Hulsken, J., Behrens, J., and Birchmeier, W.(1994a). Tumor-suppressor gene products in cell contacts: the cadherin-APC-armadillo connection. *Curr Opin Cell Biol* **6**, 711-716.
- Hulsken, J., Birchmeier, W., and Behrens, J.(1994b). E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. *J Cell Biol* **127**, 2061-2069.
- Ichii, S., Horii, A., Nakatsuru, S., Furuyama, J., Utsunomiya, J., and Nakamura, Y.(1992). Inactivation of both APC alleles in an early stage of colon adenomas in a patient with familial adenomatous polyposis (FAP). *Hum Mol Genet* **1**, 387-390.

- Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S., and Kikuchi, A.(1998). Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3 $\beta$  and  $\beta$ -catenin and promotes GSK-3 $\beta$ -dependent phosphorylation of  $\beta$ -catenin. *EMBO J* **17**, 1371-1384.
- Ilyas, M., Tomlinson, I.P., Rowan, A., Pignatelli, M., and Bodmer, W.F.(1997).  $\beta$ -catenin mutations in cell lines established from human colorectal cancers. *Proc Natl Acad Sci U S A* **94**, 10330-10334.
- Indra, A.K., Warot, X., Brocard, J., Bornert, J.M., Xiao, J.H., Chambon, P., and Metzger, D.(1999). Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res* **27**, 4324-4327.
- Ishidate, T., Matsumine, A., Toyoshima, K., and Akiyama, T.(2000). The APC-hDLG complex negatively regulates cell cycle progression from the G0/G1 to S phase. *Oncogene* **19**, 365-372.
- Itoh, K., Krupnik, V.E., and Sokol, S.Y.(1998). Axis determination in *Xenopus* involves biochemical interactions of axin, glycogen synthase kinase 3 and  $\beta$ -catenin. *Curr Biol* **8**, 591-594.
- Iwao, K., Nakamori, S., Kameyama, M., Imaoka, S., Kinoshita, M., Fukui, T., Ishiguro, S., Nakamura, Y., and Miyoshi, Y.(1998). Activation of the  $\beta$ -catenin gene by interstitial deletions involving exon 3 in primary colorectal carcinomas without adenomatous polyposis coli mutations. *Cancer Res* **58**, 1021-1026.
- Jaenisch, R.(1975). Infection of mouse blastocysts with SV40 DNA: normal development of the infected embryos and persistence of SV40-specific DNA sequences in the adult animals. *Cold Spring Harb Symp Quant Biol* **39 Pt 1**, 375-380.
- Jaiswal, A.S. and Narayan, S.(1998). Protein synthesis and transcriptional inhibitors control N-methyl-N'-nitro-N-nitrosoguanidine-induced levels of APC mRNA in a p53-dependent manner. *Int J Oncol* **13**, 733-740.
- Janson, J., Soeller, W.C., Roche, P.C., Nelson, R.T., Torchia, A.J., Kreutter, D.K., and Butler, P.C.(1996). Spontaneous diabetes mellitus in transgenic mice expressing human islet amyloid polypeptide. *Proc Natl Acad Sci U S A* **93**, 7283-7288.
- Jayaram, M.(1994). Phosphoryl transfer in Flp recombination: a template for strand transfer mechanisms. *Trends Biochem Sci* **19**, 78-82.
- Jen, J., Kim, H., Piantadosi, S., Liu, Z.F., Levitt, R.C., Sistonen, P., Kinzler, K.W., Vogelstein, B., and Hamilton, S.R.(1994). Allelic loss of chromosome 18q and prognosis in colorectal cancer [see comments]. *N Engl J Med* **331**, 213-221.
- Jin, T., Branch, D.R., Zhang, X., Qi, S., Youngson, B., and Goss, P.E.(1999). Examination of POU homeobox gene expression in human breast cancer cells. *Int J Cancer* **81**, 104-112.
- Jonsson, M., Borg, A., Nilbert, M., and Andersson, T.(2000). Involvement of adenomatous polyposis coli (APC)/ $\beta$ -catenin signalling in human breast cancer. *Eur J Cancer* **36**, 242-248.
- Joslyn, G., Richardson, D.S., White, R., and Alber, T.(1993). Dimer formation by an N-terminal coiled coil in the APC protein. *Proc Natl Acad Sci U S A* **90**, 11109-11113.
- Kamijo, T., Zindy, F., Roussel, M.F., Quelle, D.E., Downing, J.R., Ashmun, R.A., Grosveld, G., and Sherr, C.J.(1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell* **91**, 649-659.

- Kanegae, Y., Lee, G., Sato, Y., Tanaka, M., Nakai, M., Sakaki, T., Sugano, S., and Saito, I.(1995). Efficient gene activation in mammalian cells by using recombinant adenovirus expressing site-specific Cre recombinase. *Nucleic Acids Res* **23**, 3816-3821.
- Kaplan, E.L. and Meier, P.(1958). Nonparametric estimations from incomplete observations. *Journal of the American* **53**, 457-481.
- Karran, P. and Bignami, M.(1994). DNA damage tolerance, mismatch repair and genome instability. *Bioessays* **16**, 833-839.
- Keim, P. and Lark, K.G.(1990). The RecE recombination pathway mediates recombination between partially homologous DNA sequences: structural analysis of recombination products. *J Struct Biol* **104**, 97-106.
- Kellendonk, C., Troche, F., Casanova, E., Anlag, K., Opherk, C., and Schutz, G.(1999). Inducible site-specific recombination in the brain. *J Mol Biol* **285**, 175-182.
- Kellendonk, C., Tronche, F., Monaghan, A.P., Angrand, P.O., Stewart, F., and Schutz, G.(1996). Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res* **24**, 1404-1411.
- Kintner, C.(1992). Regulation of embryonic cell adhesion by the cadherin cytoplasmic domain. *Cell* **69**, 225-236.
- Kinzler, K.W., Nilbert, M.C., Su, L.K., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hedge, P., and McKechnie, D.(1991a). Identification of FAP locus genes from chromosome 5q21. *Science* **253**, 661-665.
- Kinzler, K.W., Nilbert, M.C., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hamilton, S.R., Hedge, P., and Markham, A.(1991b). Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers [see comments]. *Science* **251**, 1366-1370.
- Kishida, M., Koyama, S., Kishida, S., Matsubara, K., Nakashima, S., Higano, K., Takada, R., Takada, S., and Kikuchi, A.(1999). Axin prevents Wnt-3a-induced accumulation of beta-catenin. *Oncogene* **18**, 979-985.
- Kitani, H., Takagi, N., Atsumi, T., Kawakura, K., Imamura, K., Goto, S., Kusakabe, M., and Fukuta, K.(1996). Isolation of a germline-transmissible embryonic stem (ES) cell line from C3H/He mice. *Zoolog Sci* **13**, 865-871.
- Klingensmith, J. and Nusse, R.(1994a). Signaling by wingless in Drosophila. *Dev Biol* **166**, 396-414.
- Klingensmith, J., Nusse, R., and Perrimon, N.(1994b). The Drosophila segment polarity gene dishevelled encodes a novel protein required for response to the wingless signal. *Genes Dev* **8**, 118-130.
- Klingensmith, J., Yang, Y., Axelrod, J.D., Beier, D.R., Perrimon, N., and Sussman, D.J.(1996). Conservation of dishevelled structure and function between flies and mice: isolation and characterization of Dvl2. *Mech Dev* **58**, 15-26.
- Kolb, A.F. and Siddell, S.G.(1997). Genomic targeting of a bicistronic DNA fragment by Cre-mediated site-specific recombination. *Gene* **203**, 209-216.
- Konopleva, M., Zhao, S., Xie, Z., Segall, H., Younes, A., Claxton, D.F., Estrov, Z., Kornblau, S.M., and Andreeff, M.(1999). Apoptosis. Molecules and mechanisms. *Adv Exp Med Biol* **457**, 217-236.

- Kozak, M.(1987a). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* **15**, 8125-8148.
- Kozak, M.(1987b). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* **196**, 947-950.
- Krawczak, M. and Cooper, D.N.(1997). The Human Gene Mutation Database. *Trends Genet* **13**, 121-122.
- Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K.(1995). Inducible gene targeting in mice. *Science* **269**, 1427-1429.
- Lakso, M., Sauer, B., Mosinger, B.J., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L., and Westphal, H.(1992). Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc Natl Acad Sci U S A* **89**, 6232-6236.
- Lamlum, H., Papadopoulou, A., Ilyas, M., Rowan, A., Gillet, C., Hanby, A., Talbot, I., Bodmer, W., and Tomlinson, I.(2000). APC mutations are sufficient for the growth of early colorectal adenomas. *Proc Natl Acad Sci U S A* **97**, 2225-2228.
- Laurent, M.N., Blitz, I.L., Hashimoto, C., Rothbacher, U., and Cho, K.W.(1997). The Xenopus homeobox gene twin mediates Wnt induction of goosecoid in establishment of Spemann's organizer. *Development* **124**, 4905-4916.
- Lee, E.T.(1992). *Statistical Methods for Survival Data Analysis*, 2nd edition. Wiley Interscience,
- Lee, G. and Saito, I.(1998). Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene* **216**, 55-65.
- Lee, J.S., Ishimoto, A., and Yanagawa, S.(1999). Characterization of mouse dishevelled (Dvl) proteins in Wnt/Wingless signaling pathway. *J Biol Chem* **274**, 21464-21470.
- Lee, Y. and Park, J.(1998). A novel mutant loxP containing part of long terminal repeat of HIV-1 in spacer region: presentation of possible target site for antiviral strategy using site-specific recombinase. *Biochem Biophys Res Commun* **253**, 588-593.
- Leibovitz, A., Stinson, J.C., McCombs, W.B., McCoy, C.E., Mazur, K.C., and Mabry, N.D.(1976). Classification of human colorectal adenocarcinoma cell lines. *Cancer Res* **36**, 4562-4569.
- Lewandoski, M. and Martin, G.R.(1997). Cre-mediated chromosome loss in mice. *Nat Genet* **17**, 223-225.
- Li, L., Yuan, H., Xie, W., Mao, J., Caruso, A.M., McMahon, A., Sussman, D.J., and Wu, D.(1999). Dishevelled proteins lead to two signaling pathways. Regulation of LEF-1 and c-Jun N-terminal kinase in mammalian cells. *J Biol Chem* **274**, 129-134.
- Li, Y., Hively, W.P., and Varmus, H.E.(2000). Use of MMTV-Wnt-1 transgenic mice for studying the genetic basis of breast cancer. *Oncogene* **19**, 1002-1009.
- Lickert, H., Domon, C., Huls, G., Wehrle, C., Duluc, I., Clevers, H., Meyer, B.I., Freund, J.N., and Kemler, R.(2000). Wnt/(beta)-catenin signaling regulates the expression of the homeobox gene Cdx1 in embryonic intestine. *Development* **127**, 3805-3813.
- Lijam, N., Paylor, R., McDonald, M.P., Crawley, J.N., Deng, C.X., Herrup, K., Stevens, K.E., Maccaferri, G., McBain, C.J., Sussman, D.J., and Wynshaw-Boris, A.(1997). Social interaction and sensorimotor gating abnormalities in mice lacking Dvl1. *Cell* **90**, 895-905.

- Lipkin, M.(1997). New rodent models for studies of chemopreventive agents. *J Cell Biochem Suppl* **28-29**, 144-147.
- Liu, C., Kato, Y., Zhang, Z., Do, V.M., Yankner, B.A., and He, X.(1999). beta-Trcp couples beta-catenin phosphorylation-degradation and regulates Xenopus axis formation. *Proc Natl Acad Sci U S A* **96**, 6273-6278.
- Lorentz, O., Suh, E.R., Taylor, J.K., Boudreau, F., and Traber, P.G.(1999). CREB-binding [corrected] protein interacts with the homeodomain protein Cdx2 and enhances transcriptional activity [published erratum appears in J Biol Chem 1999 May 7;274(19):13728]. *J Biol Chem* **274**, 7196-7199.
- Lue, R.A., Marfatia, S.M., Branton, D., and Chishti, A.H.(1994). Cloning and characterization of hdlg: the human homologue of the Drosophila discs large tumor suppressor binds to protein 4.1. *Proc Natl Acad Sci U S A* **91**, 9818-9822.
- Luongo, C., Gould, K.A., Su, L.K., Kinzler, K.W., Vogelstein, B., Dietrich, W., Lander, E.S., and Moser, A.R.(1993). Mapping of multiple intestinal neoplasia (Min) to proximal chromosome 18 of the mouse. *Genomics* **15**, 3-8.
- Luongo, C., Moser, A.R., Gledhill, S., and Dove, W.F.(1994). Loss of Apc+ in intestinal adenomas from Min mice. *Cancer Res* **54**, 5947-5952.
- MacPhee, M., Chepenik, K.P., Liddell, R.A., Nelson, K.K., Siracusa, L.D., and Buchberg, A.M.(1995). The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of ApcMin-induced intestinal neoplasia. *Cell* **81**, 957-966.
- Mahmoud, N.N., Bilinski, R.T., Churchill, M.R., Edelmann, W., Kucherlapati, R., and Bertagnolli, M.M.(1999). Genotype-phenotype correlation in murine Apc mutation: differences in enterocyte migration and response to sulindac. *Cancer Res* **59**, 353-359.
- Mahmoud, N.N., Boolbol, S.K., Bilinski, R.T., Martucci, C., Chadburn, A., and Bertagnolli, M.M.(1997). Apc gene mutation is associated with a dominant-negative effect upon intestinal cell migration. *Cancer Res* **57**, 5045-5050.
- Mallo, G.V., Rechreche, H., Frigerio, J.M., Rocha, D., Zweibaum, A., Lacasa, M., Jordan, B.R., Dusetti, N.J., Dagorn, J.C., and Iovanna, J.L.(1997). Molecular cloning, sequencing and expression of the mRNA encoding human Cdx1 and Cdx2 homeobox. Down-regulation of Cdx1 and Cdx2 mRNA expression during colorectal carcinogenesis. *Int J Cancer* **74**, 35-44.
- Mallo, G.V., Soubeyran, P., Lissitzky, J.C., Andre, F., Farnarier, C., Marvaldi, J., Dagorn, J.C., and Iovanna, J.L.(1998). Expression of the Cdx1 and Cdx2 homeotic genes leads to reduced malignancy in colon cancer-derived cells. *J Biol Chem* **273**, 14030-14036.
- Mandl, M., Paffenholz, R., Friedl, W., Caspari, R., Sengteller, M., and Propping, P.(1994). Frequency of common and novel inactivating APC mutations in 202 families with familial adenomatous polyposis. *Hum Mol Genet* **3**, 181-184.
- Marcus, D.M., Rustgi, A.K., Defoe, D., Brooks, S.E., McCormick, R.S., Thompson, T.P., Edelmann, W., Kucherlapati, R., and Smith, S.(1997). Retinal pigment epithelium abnormalities in mice with adenomatous polyposis coli gene disruption. *Arch Ophthalmol* **115**, 645-650.
- Matsumine, A., Ogai, A., Senda, T., Okumura, N., Satoh, K., Baeg, G.H., Kawahara, T., Kobayashi, S., Okada, M., Toyoshima, K., and Akiyama, T.(1996). Binding of APC to the human homolog of the Drosophila discs large tumor suppressor protein [see comments]. *Science* **272**, 1020-1023.



- Maulbecker, C.C. and Gruss, P.(1993). The oncogenic potential of deregulated homeobox genes. *Cell Growth Differ* **4**, 431-441.
- McBurney, M.W., Staines, W.A., Boekelheide, K., Parry, D., Jardine, K., and Pickavance, L.(1994). Murine PGK-1 promoter drives widespread but not uniform expression in transgenic mice. *Dev Dyn* **200**, 278-293.
- McKendry, R., Hsu, S.C., Harland, R.M., and Grosschedl, R.(1997). LEF-1/TCF proteins mediate wnt-inducible transcription from the Xenopus nodal-related 3 promoter. *Dev Biol* **192**, 420-431.
- McMahon, A.P. and Moon, R.T.(1989a). Ectopic expression of the proto-oncogene int-1 in Xenopus embryos leads to duplication of the embryonic axis. *Cell* **58**, 1075-1084.
- McMahon, A.P. and Moon, R.T.(1989b). int-1--a proto-oncogene involved in cell signalling. *Development* **107 Suppl**, 161-167.
- Meselson, M.S. and Radding, C.M.(1975). A general model for genetic recombination. *Proc Natl Acad Sci U S A* **72**, 358-361.
- Mester, J., Wagenaar, E., Sluysers, M., and Nusse, R.(1987). Activation of int-1 and int-2 mammary oncogenes in hormone-dependent and -independent mammary tumors of GR mice. *J Virol* **61**, 1073-1078.
- Meyer-Leon, L., Senecoff, J.F., Bruckner, R.C., and Cox, M.M.(1984). Site-specific genetic recombination promoted by the FLP protein of the yeast 2-micron plasmid in vitro. *Cold Spring Harb Symp Quant Biol* **49**, 797-804.
- Meyer, B.I. and Gruss, P.(1993). Mouse Cdx-1 expression during gastrulation. *Development* **117**, 191-203.
- Midgley, C.A., White, S., Howitt, R., Save, V., Dunlop, M.G., Hall, P.A., Lane, D.P., Wyllie, A.H., and Bubb, V.J.(1997). APC expression in normal human tissues. *J Pathol* **181**, 426-433.
- Miller, J.R. and Moon, R.T.(1997). Analysis of the signaling activities of localization mutants of beta-catenin during axis specification in Xenopus. *J Cell Biol* **139**, 229-243.
- Miyaki, M., Konishi, M., Kikuchi-Yanoshita, R., Enomoto, M., Igari, T., Tanaka, K., Muraoka, M., Takahashi, H., Amada, Y., and Fukayama, M.(1994). Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumors. *Cancer Res* **54**, 3011-3020.
- Miyaki, M., Konishi, M., Kikuchi-Yanoshita, R., Enomoto, M., Tanaka, K., Takahashi, H., Muraoka, M., Mori, T., Konishi, F., and Iwama, T.(1993). Coexistence of somatic and germ-line mutations of APC gene in desmoid tumors from patients with familial adenomatous polyposis. *Cancer Res* **53**, 5079-5082.
- Miyashiro, I., Senda, T., Matsumine, A., Baeg, G.H., Kuroda, T., Shimano, T., Miura, S., Noda, T., Kobayashi, S., and Monden, M.(1995). Subcellular localization of the APC protein: immunoelectron microscopic study of the association of the APC protein with catenin. *Oncogene* **11**, 89-96.
- Miyoshi, Y., Iwao, K., Nagasawa, Y., Aihara, T., Sasaki, Y., Imaoka, S., Murata, M., Shimano, T., and Nakamura, Y.(1998a). Activation of the beta-catenin gene in primary hepatocellular carcinomas by somatic alterations involving exon 3. *Cancer Res* **58**, 2524-2527.



- Miyoshi, Y., Iwao, K., Nawa, G., Yoshikawa, H., Ochi, T., and Nakamura, Y.(1998b). Frequent mutations in the beta-catenin gene in desmoid tumors from patients without familial adenomatous polyposis. *Oncol Res* **10**, 591-594.
- Molenaar, M., van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., Roose, J., Destree, O., and Clevers, H.(1996). XTcf-3 transcription factor mediates beta-catenin-induced axis formation in *Xenopus* embryos. *Cell* **86**, 391-399.
- Monaco, A.P. and Larin, Z.(1994). YACs, BACs, PACs and MACs: artificial chromosomes as research tools. *Trends Biotechnol* **12**, 280-286.
- Moreau, J.F., Donaldson, D.D., Bennett, F., Witek-Giannotti, J., Clark, S.C., and Wong, G.G.(1988). Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. *Nature* **336**, 690-692.
- Morin, P.J., Sparks, A.B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K.W.(1997). Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC [see comments]. *Science* **275**, 1787-1790.
- Morrison, E.E., Askham, J., Clissold, P., Markham, A.F., and Meredith, D.M.(1997). Expression of beta-catenin and the adenomatous polyposis coli tumour suppressor protein in mouse neocortical cells in vitro. *Neurosci Lett* **235**, 129-132.
- Morrison, E.E., Wardleworth, B.N., Askham, J.M., Markham, A.F., and Meredith, D.M.(1998). EB1, a protein which interacts with the APC tumour suppressor, is associated with the microtubule cytoskeleton throughout the cell cycle. *Oncogene* **17**, 3471-3477.
- Moser, A.R., Luongo, C., Gould, K.A., McNeley, M.K., Shoemaker, A.R., and Dove, W.F.(1995). ApcMin: a mouse model for intestinal and mammary tumorigenesis. *Eur J Cancer* **31A**, 1061-1064.
- Moser, A.R., Mattes, E.M., Dove, W.F., Lindstrom, M.J., Haag, J.D., and Gould, M.N.(1993). ApcMin, a mutation in the murine Apc gene, predisposes to mammary carcinomas and focal alveolar hyperplasias. *Proc Natl Acad Sci U S A* **90**, 8977-8981.
- Moser, A.R., Pitot, H.C., and Dove, W.F.(1990). A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* **247**, 322-324.
- Moser, A.R., Shoemaker, A.R., Connelly, C.S., Clipson, L., Gould, K.A., Luongo, C., Dove, W.F., Siggers, P.H., and Gardner, R.L.(1995). Homozygosity for the Min allele of Apc results in disruption of mouse development prior to gastrulation. *Dev Dyn* **203**, 422-433.
- Munemitsu, S., Albert, I., Rubinfeld, B., Polakis, P., Kanai, Y., Ochiai, A., Shibata, T., Oyama, T., Ushijima, S., Akimoto, S., and Hirohashi, S.(1996). Deletion of an amino-terminal sequence beta-catenin in vivo and promotes hyperphosphorylation of the adenomatous polyposis coli tumor suppressor protein c-erbB-2 gene product directly associates with beta-catenin and plakoglobin. *Mol Cell Biol* **16**, 4088-4094.
- Munemitsu, S., Souza, B., Muller, O., Albert, I., Rubinfeld, B., and Polakis, P.(1994). The APC gene product associates with microtubules in vivo and promotes their assembly in vitro. *Cancer Res* **54**, 3676-3681.
- Muyrers, J.P., Zhang, Y., Testa, G., and Stewart, A.F.(1999). Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res* **27**, 1555-1557.
- Nagase, H. and Nakamura, Y.(1993). Mutations of the APC (adenomatous polyposis coli) gene. *Hum Mutat* **2**, 425-434.

- Nakamura, T., Hamada, F., Ishidate, T., Anai, K., Kawahara, K., Toyoshima, K., and Akiyama, T.(1998). Axin, an inhibitor of the Wnt signalling pathway, interacts with beta-catenin, GSK-3beta and APC and reduces the beta-catenin level. *Genes Cells* **3**, 395-403.
- Narayan, S. and Jaiswal, A.S.(1997). Activation of adenomatous polyposis coli (APC) gene expression by the DNA-alkylating agent N-methyl-N'-nitro-N-nitrosoguanidine requires p53. *J Biol Chem* **272**, 30619-30622.
- Nathke, I.S., Adams, C.L., Polakis, P., Sellin, J.H., and Nelson, W.J.(1996). The adenomatous polyposis coli tumor suppressor protein localizes to plasma membrane sites involved in active cell migration. *J Cell Biol* **134**, 165-179.
- Nathke, I.S., Hinck, L., Swedlow, J.R., Papkoff, J., and Nelson, W.J.(1994). Defining interactions and distributions of cadherin and catenin complexes in polarized epithelial cells. *J Cell Biol* **125**, 1341-1352.
- Neufeld, K.L. and White, R.L.(1997). Nuclear and cytoplasmic localizations of the adenomatous polyposis coli protein. *Proc Natl Acad Sci U S A* **94**, 3034-3039.
- Neugut, A.I., Jacobson, J.S., Suh, S., Mukherjee, R., and Arber, N.(1998). The epidemiology of cancer of the small bowel. *Cancer Epidemiol Biomarkers Prev* **7**, 243-251.
- Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S., and Hedge, P.(1991). Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **253**, 665-669.
- Noordermeer, J., Klingensmith, J., Perrimon, N., and Nusse, R.(1994). dishevelled and armadillo act in the wingless signalling pathway in *Drosophila*. *Nature* **367**, 80-83.
- Nusse, R.(1997). A versatile transcriptional effector of Wingless signaling. *Cell* **89**, 321-323.
- Nusse, R.(1999). WNT targets. Repression and activation. *Trends Genet* **15**, 1-3.
- Nusse, R., Samos, C.H., Brink, M., Willert, K., Cadigan, K.M., Wodarz, A., Fish, M., and Rulifson, E.(1997). Cell culture and whole animal approaches to understanding signaling by Wnt proteins in *Drosophila*. *Cold Spring Harb Symp Quant Biol* **62**, 185-190.
- Nusse, R., Theunissen, H., Wagenaar, E., Rijsewijk, F., Gennissen, A., Otte, A., Schuurin, E., and van Ooyen, A.(1990). The Wnt-1 (int-1) oncogene promoter and its mechanism of activation by insertion of proviral DNA of the mouse mammary tumor virus. *Mol Cell Biol* **10**, 4170-4179.
- Nusse, R. and Varmus, H.E.(1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **31**, 99-109.
- Nusslein-Volhard, C.(1991). Determination of the embryonic axes of *Drosophila*. *Dev Suppl* **1**, 1-10.
- Nusslein-Volhard, C. and Roth, S.(1989). Axis determination in insect embryos. *Ciba Found Symp* **144**, 37-38.
- Orban, P.C., Chui, D., and Marth, J.D.(1992). Tissue- and site-specific DNA recombination in transgenic mice. *Proc Natl Acad Sci U S A* **89**, 6861-6865.
- Orford, K., Crockett, C., Jensen, J.P., Weissman, A.M., and Byers, S.W.(1997). Serine phosphorylation-regulated ubiquitination and degradation of beta-catenin. *J Biol Chem* **272**, 24735-24738.

- Oshima, M., Sugiyama, H., Kitagawa, K., and Taketo, M.(1993). APC gene messenger RNA: novel isoforms that lack exon 7. *Cancer Res* **53**, 5589-5591.
- Ozawa, M., Terada, H., and Pedraza, C.(1995). The fourth armadillo repeat of plakoglobin (gamma-catenin) is required for its high affinity binding to the cytoplasmic domains of E-cadherin and desmosomal cadherin Dsg2, and the tumor suppressor APC protein. *J Biochem (Tokyo)* **118**, 1077-1082.
- Palmiter, R.D. and Brinster, R.L.(1986). Germ-line transformation of mice. *Annu Rev Genet* **20**, 465-499.
- Park, M., Wu, X., Golden, K., Axelrod, J.D., and Bodmer, R.(1996). The wingless signaling pathway is directly involved in Drosophila heart development. *Dev Biol* **177**, 104-116.
- Parker, S.L., Tong, T., Bolden, S., and Wingo, P.A.(1996). Cancer statistics, 1996 [see comments]. *CA Cancer J Clin* **46**, 5-27.
- Pfarr, D.S., Rieser, L.A., Woychik, R.P., Rottman, F.M., Rosenberg, M., and Reff, M.E.(1986). Differential effects of polyadenylation regions on gene expression in mammalian cells. *DNA* **5**, 115-122.
- Picard, D., Khursheed, B., Garabedian, M.J., Fortin, M.G., Lindquist, S., and Yamamoto, K.R.(1990). Reduced levels of hsp90 compromise steroid receptor action in vivo. *Nature* **348**, 166-168.
- Polakis, P.(1995). Mutations in the APC gene and their implications for protein structure and function. *Curr Opin Genet Dev* **5**, 66-71.
- Polakis, P.(1997). The adenomatous polyposis coli (APC) tumor suppressor. *Biochim Biophys Acta* **1332**, F127-F147.
- Polakis, P.(1999). The oncogenic activation of beta-catenin. *Curr Opin Genet Dev* **9**, 15-21.
- Ponting, C.P., Phillips, C., Davies, K.E., and Blake, D.J.(1997). PDZ domains: targeting signalling molecules to sub-membranous sites. *Bioessays* **19**, 469-479.
- Potten, C.S. and Booth, C.(1997). The role of radiation-induced and spontaneous apoptosis in the homeostasis of the gastrointestinal epithelium: a brief review. *Comp Biochem Physiol B Biochem Mol Biol* **118**, 473-478.
- Povey, A.C., Hall, C.N., Badawi, A.F., Cooper, D.P., and O'Connor, P.J.(2000). Elevated levels of the pro-carcinogenic adduct, O(6)-methylguanine, in normal DNA from the cancer prone regions of the large bowel. *Gut* **47**, 362-365.
- Powell, S.M., Zilz, N., Beazer-Barclay, Y., Bryan, T.M., Hamilton, S.R., Thibodeau, S.N., Vogelstein, B., and Kinzler, K.W.(1992). APC mutations occur early during colorectal tumorigenesis. *Nature* **359**, 235-237.
- Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J.(1992). Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111**, 229-233.
- Pravtcheva, D.D., Adra, C.N., and Ruddle, F.H.(1991). Timing of paternal Pgk-1 expression in embryos of transgenic mice. *Development* **111**, 1109-1120.
- Pyles, R.B., Santoro, I.M., Groden, J., and Parysek, L.M.(1998). Novel protein isoforms of the APC tumor suppressor in neural tissue. *Oncogene* **16**, 77-82.

- Ramsay, R.G., Friend, A., Vizantios, Y., Freeman, R., Sicurella, C., Hammett, F., Armes, J., and Venter, D.(2000). Cyclooxygenase-2, a colorectal cancer nonsteroidal anti-inflammatory drug target, is regulated by c-MYB. *Cancer Res* **60**, 1805-1809.
- Ramsay, R.G., Thompson, M.A., Hayman, J.A., Reid, G., Gonda, T.J., and Whitehead, R.H.(1992). Myb expression is higher in malignant human colonic carcinoma and premalignant adenomatous polyps than in normal mucosa. *Cell Growth Differ* **3**, 723-730.
- Ray, M.K., Fagan, S.P., Moldovan, S., DeMayo, F.J., and Brunicardi, F.C.(1999). Beta cell-specific ablation of target gene using Cre-loxP system in transgenic mice. *J Surg Res* **84**, 199-203.
- Rocheleau, C.E., Downs, W.D., Lin, R., Wittmann, C., Bei, Y., Cha, Y.H., Ali, M., Priess, J.R., and Mello, C.C.(1997). Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos [see comments]. *Cell* **90**, 707-716.
- Rohlmann, A., Gotthardt, M., Hammer, R.E., and Herz, J.(1998). Inducible inactivation of hepatic LRP gene by cre-mediated recombination confirms role of LRP in clearance of chylomicron remnants. *J Clin Invest* **101**, 689-695.
- Roose, J., Molenaar, M., Peterson, J., Hurenkamp, J., Brantjes, H., Moerer, P., van de Wetering, M., Destree, O., and Clevers, H.(1998). The *Xenopus* Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature* **395**, 608-612.
- Rosen, A. and Casciola-Rosen, L.(1997). Macromolecular substrates for the ICE-like proteases during apoptosis. *J Cell Biochem* **64**, 50-54.
- Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S., and Polakis, P.(1996). Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly [see comments]. *Science* **272**, 1023-1026.
- Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, E., and Polakis, P.(1997). Stabilization of beta-catenin by genetic defects in melanoma cell lines [see comments]. *Science* **275**, 1790-1792.
- Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S., and Polakis, P.(1995). The APC protein and E-cadherin form similar but independent complexes with alpha-catenin, beta-catenin, and plakoglobin. *J Biol Chem* **270**, 5549-5555.
- Ruel, L., Stambolic, V., Ali, A., Manoukian, A.S., and Woodgett, J.R.(1999). Regulation of the protein kinase activity of Shaggy(Zeste-white3) by components of the wingless pathway in *Drosophila* cells and embryos. *J Biol Chem* **274**, 21790-21796.
- Rulifson, E.J., Micchelli, C.A., Axelrod, J.D., Perrimon, N., and Blair, S.S.(1996). wingless refines its own expression domain on the *Drosophila* wing margin. *Nature* **384**, 72-74.
- Saam, J.R. and Gordon, J.I.(1999). Inducible gene knockouts in the small intestinal and colonic epithelium. *J Biol Chem* **274**, 38071-38082.
- Sakai, K., Mitani, K., and Miyazaki, J.(1995). Efficient regulation of gene expression by adenovirus vector-mediated delivery of the CRE recombinase. *Biochem Biophys Res Commun* **217**, 393-401.
- Sakanaka, C., Leong, P., Xu, L., Harrison, S.D., and Williams, L.T.(1999). Casein kinase epsilon in the wnt pathway: regulation of beta-catenin function. *Proc Natl Acad Sci U S A* **96**, 12548-12552.

- Sakanaka, C., Weiss, J.B., and Williams, L.T.(1998). Bridging of beta-catenin and glycogen synthase kinase-3 $\beta$  by axin and inhibition of beta-catenin-mediated transcription. *Proc Natl Acad Sci U S A* **95**, 3020-3023.
- Sakanaka, C. and Williams, L.T.(1999). Functional domains of axin. Importance of the C terminus as an oligomerization domain. *J Biol Chem* **274**, 14090-14093.
- Salovaara, R., Loukola, A., Kristo, P., Kaariainen, H., Ahtola, H., Eskelinen, M., Harkonen, N., Julkunen, R., Kangas, E., Ojala, S., Tulikoura, J., Valkamo, E., Jarvinen, H., Mecklin, J.P., Aaltonen, L.A., and de la Chapelle, A.(2000). Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol* **18**, 2193-2200.
- Sambrook, J, Fritsch, EF, and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. (1989).(Generic)
- Samowitz, W.S., Thliveris, A., Spirio, L.N., and White, R.(1995). Alternatively spliced adenomatous polyposis coli (APC) gene transcripts that delete exons mutated in attenuated APC. *Cancer Res* **55**, 3732-3734.
- Sanger, F., Nicklen, S., and Coulson, A.R.(1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.
- Santoro, I.M. and Groden, J.(1997). Alternative splicing of the APC gene and its association with terminal differentiation. *Cancer Res* **57**, 488-494.
- Santos, J., Herranz, M., Perez, d.C., Pellicer, A., and Fernandez-Piqueras, J.(1998). A new candidate site for a tumor suppressor gene involved in mouse thymic lymphomagenesis is located on the distal part of chromosome 4. *Oncogene* **17**, 925-929.
- Sato, N., Wang, S., Li, L., Okabe, K., Hashimoto, M., Yaginuma, H., Mikoshiba, K., Uchiyama, Y., Uetsuki, T., Yoshikawa, K., Milligan, C.E., and Oppenheim, R.W.(1998). A novel strategy for introducing exogenous bcl-2 into neuronal cells: the Cre/loxP system-mediated activation of bcl-2 for preventing programmed cell death using recombinant adenoviruses. *Mol Cell Neurosci* **12**, 65-78.
- Sauer, B.(1987). Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* **7**, 2087-2096.
- Sauer, B.(1998). Inducible gene targeting in mice using the Cre/lox system. *Methods* **14**, 381-392.
- Sauer, B. and Henderson, N.(1990). Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. *New Biol* **2**, 441-449.
- Schwartz, C.J. and Sadowski, P.D.(1990). FLP protein of 2  $\mu$  circle plasmid of yeast induces multiple bends in the FLP recognition target site. *J Mol Biol* **216**, 289-298.
- Seeling, J.M., Miller, J.R., Gil, R., Moon, R.T., White, R., and Virshup, D.M.(1999). Regulation of beta-catenin signaling by the B56 subunit of protein phosphatase 2A. *Science* **283**, 2089-2091.
- Segev, N. and Cohen, G.(1981). Control of circularization of bacteriophage P1 DNA in *Escherichia coli*. *Virology* **114**, 333-342.
- Selbert, S., Bentley, D.J., Melton, D.W., Rannie, D., Lourenco, P., Watson, C.J., and Clarke, A.R.(1998). Efficient BLG-Cre mediated gene deletion in the mammary gland. *Transgenic Res* **7**, 387-396.



- Sheldahl, L.C., Park, M., Malbon, C.C., and Moon, R.T.(1999). Protein kinase C is differentially stimulated by Wnt and Frizzled homologs in a G-protein-dependent manner. *Curr Biol* **9**, 695-698.
- Shibata, H., Toyama, K., Shioya, H., Ito, M., Hirota, M., Hasegawa, S., Matsumoto, H., Takano, H., Akiyama, T., Toyoshima, K., Kanamaru, R., Kanegae, Y., Saito, I., Nakamura, Y., Shiba, K., and Noda, T.(1997). Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science* **278**, 120-123.
- Shintani, Y., Yotsuyanagi, H., Moriya, K., Fujie, H., Tsutsumi, T., Kanegae, Y., Kimura, S., Saito, I., and Koike, K.(1999). Induction of apoptosis after switch-on of the hepatitis B virus X gene mediated by the Cre/loxP recombination system. *J Gen Virol* **80** ( Pt 12), 3257-3265.
- Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M.(1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**, 8794-8797.
- Shtutman, M., Zhurinsky, J., Simcha, I., Albanese, C., D'Amico, M., Pestell, R., and Ben-Ze'ev, A.(1999). The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. *Proc Natl Acad Sci U S A* **96**, 5522-5527.
- Siegfried, E., Chou, T.B., and Perrimon, N.(1992). wingless signaling acts through zeste-white 3, the *Drosophila* homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. *Cell* **71**, 1167-1179.
- Silberg, D.G., Furth, E.E., Taylor, J.K., Schuck, T., Chiou, T., and Traber, P.G.(1997). CDX1 protein expression in normal, metaplastic, and neoplastic human alimentary tract epithelium [see comments]. *Gastroenterology* **113**, 478-486.
- Slusarski, D.C., Corces, V.G., and Moon, R.T.(1997). Interaction of Wnt and a Frizzled homologue triggers G-protein-linked phosphatidylinositol signalling. *Nature* **390**, 410-413.
- Smith, A.G. and Hooper, M.L.(1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev Biol* **121**, 1-9.
- Smith, K.J., Johnson, K.A., Bryan, T.M., Hill, D.E., Markowitz, S., Willson, J.K., Paraskeva, C., Petersen, G.M., Hamilton, S.R., and Vogelstein, B.(1993). The APC gene product in normal and tumor cells. *Proc Natl Acad Sci U S A* **90**, 2846-2850.
- Smith, K.J., Levy, D.B., Maupin, P., Pollard, T.D., Vogelstein, B., and Kinzler, K.W.(1994). Wild-type but not mutant APC associates with the microtubule cytoskeleton. *Cancer Res* **54**, 3672-3675.
- Smits, R., Kartheuser, A., Jagmohan-Changur, S., Leblanc, V., Breukel, C., de Vries, A., van Kranen, H., van Krieken, J.H., Williamson, S., Edelmann, W., Kucherlapati, R., KhanPM, and Fodde, R.(1997). Loss of Apc and the entire chromosome 18 but absence of mutations at the Ras and Tp53 genes in intestinal tumors from Apc1638N, a mouse model for Apc-driven carcinogenesis. *Carcinogenesis* **18**, 321-327.
- Smits, R., van der Houven van Oordt, Luz, A., Zurcher, C., Jagmohan-Changur, S., Breukel, C., Khan, P.M., and Fodde, R.(1998). Apc1638N: a mouse model for familial adenomatous polyposis-associated desmoid tumors and cutaneous cysts. *Gastroenterology* **114**, 275-283.



- Songyang, Z., Fanning, A.S., Fu, C., Xu, J., Marfatia, S.M., Chishti, A.H., Crompton, A., Chan, A.C., Anderson, J.M., and Cantley, L.C.(1997). Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275**, 73-77.
- Soravia, C., Bapat, B., and Cohen, Z.(1997). Familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC): a review of clinical, genetic and therapeutic aspects. *Schweiz Med Wochenschr* **127**, 682-690.
- Southern, P.J. and Berg, P.(1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J Mol Appl Genet* **1**, 327-341.
- Sparks, A.B., Morin, P.J., Vogelstein, B., and Kinzler, K.W.(1998). Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. *Cancer Res* **58**, 1130-1134.
- Spirio, L., Olschwang, S., Groden, J., Robertson, M., Samowitz, W., Joslyn, G., Gelbert, L., Thliveris, A., Carlson, M., and Otterud, B.(1993). Alleles of the APC gene: an attenuated form of familial polyposis. *Cell* **75**, 951-957.
- Spirio, L., Otterud, B., Stauffer, D., Lynch, H., Lynch, P., Watson, P., Lanspa, S., Smyrk, T., Cavalieri, J., and Howard, L.(1992). Linkage of a variant or attenuated form of adenomatous polyposis coli to the adenomatous polyposis coli (APC) locus. *Am J Hum Genet* **51**, 92-100.
- St-Onge, L., Furth, P.A., and Gruss, P.(1996). Temporal control of the Cre recombinase in transgenic mice by a tetracycline responsive promoter. *Nucleic Acids Res* **24**, 3875-3877.
- Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A., and Melton, D.W.(1994). Use of double-replacement gene targeting to replace the murine alpha-lactalbumin gene with its human counterpart in embryonic stem cells and mice. *Mol Cell Biol* **14**, 1009-1016.
- Stec, D.E., Davisson, R.L., Haskell, R.E., Davidson, B.L., and Sigmund, C.D.(1999). Efficient liver-specific deletion of a floxed human angiotensinogen transgene by adenoviral delivery of Cre recombinase in vivo. *J Biol Chem* **274**, 21285-21290.
- Stewart, D.B. and Nelson, W.J.(1997). Identification of four distinct pools of catenins in mammalian cells and transformation-dependent changes in catenin distributions among these pools. *J Biol Chem* **272**, 29652-29662.
- Strand, M., Prolla, T.A., Liskay, R.M., and Petes, T.D.(1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair [see comments] [published erratum appears in Nature 1994 Apr 7;368(6471);569]. *Nature* **365**, 274-276.
- Sturm, K., Lafferty, M., and Tam, P.P.(1999). Pgk1 and Hpvt gene activity in the peri-implantation mouse embryo is influenced by the parental origin of the X-chromosome. *Int J Dev Biol* **43**, 69-73.
- Su, L.K., Barnes, C.J., Yao, W., Qi, Y., Lynch, P.M., and Steinbach, G.(2000). Inactivation of germline mutant APC alleles by attenuated somatic mutations: a molecular genetic mechanism for attenuated familial adenomatous polyposis. *Am J Hum Genet* **67**, 582-590.
- Su, L.K., Burrell, M., Hill, D.E., Gyuris, J., Brent, R., Wiltshire, R., Trent, J., Vogelstein, B., and Kinzler, K.W.(1995). APC binds to the novel protein EB1. *Cancer Res* **55**, 2972-2977.
- Su, L.K., Johnson, K.A., Smith, K.J., Hill, D.E., Vogelstein, B., and Kinzler, K.W.(1993). Association between wild type and mutant APC gene products. *Cancer Res* **53**, 2728-2731.

- Su, L.K., Kinzler, K.W., Vogelstein, B., Preisinger, A.C., Moser, A.R., Luongo, C., Gould, K.A., and Dove, W.F.(1992). Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene [published erratum appears in *Science* 1992 May 22;256(5060):1114]. *Science* **256**, 668-670.
- Su, L.K., Vogelstein, B., and Kinzler, K.W.(1993). Association of the APC tumor suppressor protein with catenins. *Science* **262**, 1734-1737.
- Subramanian, V., Meyer, B., and Evans, G.S.(1998). The murine Cdx1 gene product localises to the proliferative compartment in the developing and regenerating intestinal epithelium. *Differentiation* **64**, 11-18.
- Szabo, C.I. and King, M.C.(1995). Inherited breast and ovarian cancer. *Hum Mol Genet* **4 Spec No**, 1811-1817.
- Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W.(1983). The double-strand-break repair model for recombination. *Cell* **33**, 25-35.
- Tao, Y.S., Edwards, R.A., Tubb, B., Wang, S., Bryan, J., and McCrea, P.D.(1996). beta-Catenin associates with the actin-bundling protein fascin in a noncadherin complex. *J Cell Biol* **134**, 1271-1281.
- Terry, R.W., Kwee, L., Baldwin, H.S., and Labow, M.A.(1997). Cre-mediated generation of a VCAM-1 null allele in transgenic mice. *Transgenic Res* **6**, 349-356.
- Theisen, H., Purcell, J., Bennett, M., Kansagara, D., Syed, A., and Marsh, J.L.(1994). dishevelled is required during wingless signaling to establish both cell polarity and cell identity. *Development* **120**, 347-360.
- Thliveris, A., Samowitz, W., Matsunami, N., Groden, J., and White, R.(1994). Demonstration of promoter activity and alternative splicing in the region 5' to exon 1 of the APC gene. *Cancer Res* **54**, 2991-2995.
- Thomas, G.M., Frame, S., Goedert, M., Nathke, I., Polakis, P., and Cohen, P.(1999). A GSK3-binding peptide from FRAT1 selectively inhibits the GSK3-catalysed phosphorylation of axin and beta-catenin. *FEBS Lett* **458**, 247-251.
- Thompson, A.M., Morris, R.G., Wallace, M., Wyllie, A.H., Steel, C.M., and Carter, D.C.(1993). Allele loss from 5q21 (APC/MCC) and 18q21 (DCC) and DCC mRNA expression in breast cancer. *Br J Cancer* **68**, 64-68.
- Thompson, M.A., Rosenthal, M.A., Ellis, S.L., Friend, A.J., Zorbas, M.I., Whitehead, R.H., and Ramsay, R.G.(1998). c-Myb down-regulation is associated with human colon cell differentiation, apoptosis, and decreased Bcl-2 expression. *Cancer Res* **58**, 5168-5175.
- Toft, N.J., Winton, D.J., Kelly, J., Howard, L.A., Dekker, M., te, R., Arends, M.J., Wyllie, A.H., Margison, G.P., and Clarke, A.R.(1999). Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. *Proc Natl Acad Sci U S A* **96**, 3911-3915.
- Trzepacz, C., Lowy, A.M., Kordich, J.J., and Groden, J.(1997). Phosphorylation of the tumor suppressor adenomatous polyposis coli (APC) by the cyclin-dependent kinase p34. *J Biol Chem* **272**, 21681-21684.
- Tsukamoto, A.S., Grosschedl, R., Guzman, R.C., Parslow, T., and Varmus, H.E.(1988). Expression of the int-1 gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. *Cell* **55**, 619-625.
- Untawale, S. and Blick, M.(1988). Oncogene expression in adenocarcinomas of the colon and in colon tumor-derived cell lines. *Anticancer Res* **8**, 1-7.

- van de Wetering, M., Cavallo, R., Dooijes, D., van Beest, M., van Es, J., Loureiro, J., Ypma, A., Hursh, D., Jones, T., Bejsovec, A., Peifer, M., Mortin, M., and Clevers, H.(1997). Armadillo coactivates transcription driven by the product of the *Drosophila* segment polarity gene dTCF. *Cell* **88**, 789-799.
- Vasicek, T.J., Zeng, L., Guan, X.J., Zhang, T., Costantini, F., and Tilghman, S.M.(1997). Two dominant mutations in the mouse fused gene are the result of transposon insertions. *Genetics* **147**, 777-786.
- Vasioukhin, V., Degenstein, L., Wise, B., and Fuchs, E.(1999). The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc Natl Acad Sci U S A* **96**, 8551-8556.
- Venkatesh, T.V., Park, M., Ocorr, K., Nemaceck, J., Golden, K., Wemple, M., and Bodmer, R.(2000). Cardiac enhancer activity of the homeobox gene tinman depends on CREB consensus binding sites in *Drosophila*. *Genesis* **26**, 55-66.
- Vider, B.Z., Zimmer, A., Chastre, E., Gespach, C., Halperin, M., Mashiah, P., Yaniv, A., and Gazit, A.(2000). Deregulated expression of homeobox-containing genes, HOXB6, B8, C8, C9, and Cdx-1, in human colon cancer cell lines. *Biochem Biophys Res Commun* **272**, 513-518.
- Vinson, C.R. and Adler, P.N.(1987). Directional non-cell autonomy and the transmission of polarity information by the frizzled gene of *Drosophila*. *Nature* **329**, 549-551.
- Vinson, C.R., Conover, S., and Adler, P.N.(1989). A *Drosophila* tissue polarity locus encodes a protein containing seven potential transmembrane domains. *Nature* **338**, 263-264.
- Vleminckx, K., Kemler, R., and Hecht, A.(1999). The C-terminal transactivation domain of beta-catenin is necessary and sufficient for signaling by the LEF-1/beta-catenin complex in *Xenopus laevis*. *Mech Dev* **81**, 65-74.
- Vleminckx, K., Wong, E., Guger, K., Rubinfeld, B., Polakis, P., and Gumbiner, B.M.(1997). Adenomatous polyposis coli tumor suppressor protein has signaling activity in *Xenopus laevis* embryos resulting in the induction of an ectopic dorsoanterior axis. *J Cell Biol* **136**, 411-420.
- Vogelstein, B. and Kinzler, K.W.(1993). The multistep nature of cancer. *Trends Genet* **9**, 138-141.
- Volkert, F.C. and Broach, J.R.(1986a). Site-specific recombination promotes plasmid amplification in yeast. *Cell* **46**, 541-550.
- Volkert, F.C., Wu, L.C., Fisher, P.A., and Broach, J.R.(1986b). Survival strategies of the yeast plasmid two-micron circle. *Basic Life Sci* **40**, 375-396.
- Wagner, K.U., Wall, R.J., St-Onge, L., Gruss, P., Wynshaw-Boris, A., Garrett, L., Li, M., Furth, P.A., and Hennighausen, L.(1997). Cre-mediated gene deletion in the mammary gland. *Nucleic Acids Res* **25**, 4323-4330.
- Wagner, U., Utton, M., Gallo, J.M., and Miller, C.C.(1996). Cellular phosphorylation of tau by GSK-3 beta influences tau binding to microtubules and microtubule organisation. *J Cell Sci* **109** ( Pt 6), 1537-1543.
- Wakita, T., Taya, C., Katsume, A., Kato, J., Yonekawa, H., Kanegae, Y., Saito, I., Hayashi, Y., Koike, M., and Kohara, M.(1998). Efficient conditional transgene expression in hepatitis C virus cDNA transgenic mice mediated by the Cre/loxP system. *J Biol Chem* **273**, 9001-9006.

- Wallis, Y.L., Macdonald, F., Hulten, M., Morton, J.E., McKeown, C.M., Neoptolemos, J.P., Keighley, M., and Morton, D.G.(1994). Genotype-phenotype correlation between position of constitutional APC gene mutation and CHRPE expression in familial adenomatous polyposis. *Hum Genet* **94**, 543-548.
- Wallis, Y.L., Morton, D.G., McKeown, C.M., and Macdonald, F.(1999). Molecular analysis of the APC gene in 205 families: extended genotype-phenotype correlations in FAP and evidence for the role of APC amino acid changes in colorectal cancer predisposition. *J Med Genet* **36**, 14-20.
- Waltzer, L. and Bienz, M.(1998). Drosophila CBP represses the transcription factor TCF to antagonize Wingless signalling. *Nature* **395**, 521-525.
- Wang, S. and Hazelrigg, T.(1994). Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis. *Nature* **369**, 400-403.
- Wang, Y., Krushel, L.A., and Edelman, G.M.(1996). Targeted DNA recombination in vivo using an adenovirus carrying the cre recombinase gene. *Proc Natl Acad Sci U S A* **93**, 3932-3936.
- Webb, S.J., Nicholson, D., Bubb, V.J., and Wyllie, A.H.(1999). Caspase-mediated cleavage of APC results in an amino-terminal fragment with an intact armadillo repeat domain. *FASEB J* **13**, 339-346.
- Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M.(1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**, 684-687.
- Wilson, C.L., Heppner, K.J., Labosky, P.A., Hogan, B.L., and Matrisian, L.M.(1997). Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. *Proc Natl Acad Sci U S A* **94**, 1402-1407.
- Wodarz, A. and Nusse, R.(1998). Mechanisms of Wnt signaling in development. *Annu Rev Cell Dev Biol* **14**, 59-88.
- Wong, M.H., Hermiston, M.L., Syder, A.J., and Gordon, J.I.(1996). Forced expression of the tumor suppressor adenomatosis polyposis coli protein induces disordered cell migration in the intestinal epithelium. *Proc Natl Acad Sci U S A* **93**, 9588-9593.
- Woo, M., Hakem, R., Soengas, M.S., Duncan, G.S., Shahinian, A., Kagi, D., Hakem, A., McCurrach, M., Khoo, W., Kaufman, S.A., Senaldi, G., Howard, T., Lowe, S.W., and Mak, T.W.(1998). Essential contribution of caspase 3/CPP32 to apoptosis and its associated nuclear changes. *Genes Dev* **12**, 806-819.
- Woodhouse, E., Hersperger, E., and Shearn, A.(1998). Growth, metastasis, and invasiveness of Drosophila tumors caused by mutations in specific tumor suppressor genes. *Dev Genes Evol* **207**, 542-550.
- Wu, X., Golden, K., and Bodmer, R.(1995). Heart development in Drosophila requires the segment polarity gene wingless. *Dev Biol* **169**, 619-628.
- Yagi, O.K., Akiyama, Y., and Yuasa, Y.(1999). Genomic structure and alterations of homeobox gene CDX2 in colorectal carcinomas. *Br J Cancer* **79**, 440-444.
- Yamamoto, H., Kishida, S., Kishida, M., Ikeda, S., Takada, S., and Kikuchi, A.(1999). Phosphorylation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3 $\beta$  regulates its stability. *J Biol Chem* **274**, 10681-10684.

- Yamamoto, H., Kishida, S., Uochi, T., Ikeda, S., Koyama, S., Asashima, M., and Kikuchi, A.(1998). Axil, a member of the Axin family, interacts with both glycogen synthase kinase 3 $\beta$  and beta-catenin and inhibits axis formation of *Xenopus* embryos. *Mol Cell Biol* **18**, 2867-2875.
- Yamamura, K., Miyazaki, T., Uno, M., and Miyazaki, J.(1992). Transgenic mouse as a tool for the study of autoimmune disease: insulin-dependent diabetes mellitus. *Int J Immunopharmacol* **14**, 451-455.
- Yanagawa, S., van Leeuwen, F., Wodarz, A., Klingensmith, J., and Nusse, R.(1995). The dishevelled protein is modified by wingless signaling in *Drosophila*. *Genes Dev* **9**, 1087-1097.
- Yang-Snyder, J., Miller, J.R., Brown, J.D., Lai, C.J., and Moon, R.T.(1996). A frizzled homolog functions in a vertebrate Wnt signaling pathway. *Curr Biol* **6**, 1302-1306.
- Yang, K., Edelmann, W., Fan, K., Lau, K., Kolli, V.R., Fodde, R., Khan, P.M., Kucherlapati, R., and Lipkin, M.(1997). A mouse model of human familial adenomatous polyposis. *J Exp Zool* **277**, 245-254.
- Yang, K., Edelmann, W., Fan, K., Lau, K., Leung, D., Newmark, H., Kucherlapati, R., and Lipkin, M.(1998). Dietary modulation of carcinoma development in a mouse model for human familial adenomatous polyposis. *Cancer Res* **58**, 5713-5717.
- Yang, Y., Lijam, N., Sussman, D.J., and Tsang, M.(1996). Genomic organization of mouse Dishevelled genes. *Gene* **180**, 121-123.
- Yang, Y., Nunes, F.A., Berencsi, K., Furth, E.E., Gonczol, E., and Wilson, J.M.(1994). Cellular immunity to viral antigens limits E1-deleted adenoviruses for gene therapy. *Proc Natl Acad Sci U S A* **91**, 4407-4411.
- Yokoya, F., Imamoto, N., Tachibana, T., and Yoneda, Y.(1999). beta-catenin can be transported into the nucleus in a Ran-unassisted manner. *Mol Biol Cell* **10**, 1119-1131.
- Yost, C., Torres, M., Miller, J.R., Huang, E., Kimelman, D., and Moon, R.T.(1996). The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in *Xenopus* embryos by glycogen synthase kinase 3. *Genes Dev* **10**, 1443-1454.
- Zeng, L., Fagotto, F., Zhang, T., Hsu, W., Vasicek, T.J., Perry, W.L., Lee, J.J., Tilghman, S.M., Gumbiner, B.M., and Costantini, F.(1997). The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell* **90**, 181-192.
- Zhang, Y., Riesterer, C., Ayral, A.M., Sablitzky, F., Littlewood, T.D., and Reth, M.(1996). Inducible site-directed recombination in mouse embryonic stem cells. *Nucleic Acids Res* **24**, 543-548.
- Zheng, L., Zhang, J., and Carthew, R.W.(1995). frizzled regulates mirror-symmetric pattern formation in the *Drosophila* eye. *Development* **121**, 3045-3055.
- Zhou, L., Dey, C.R., Wert, S.E., DuVall, M.D., Frizzell, R.A., and Whitsett, J.A.(1994). Correction of lethal intestinal defect in a mouse model of cystic fibrosis by human CFTR. *Science* **266**, 1705-1708.
- Zhou, Z., Wang, J., Han, X., Zhou, J., and Linder, S.(1998). Up-regulation of human secreted frizzled homolog in apoptosis and its down-regulation in breast tumors. *Int J Cancer* **78**, 95-99.

# **Appendix 1**

## **Solutions**



**40% Acrylamide/Bisacrylamide**

380g            Acrylamide (Kodak 5521)

20g            N,N-Methylene-bisacrylamide (Kodak 8383)

Dissolve in approx. 800ml of deionised water and adjust to a final volume of 1L with deionised water (store at 4°C).

**10x Agarose gel loading dye**

1.5g            Ficoll (Sigma F-2637)

0.02g          Bromophenol blue (Sigma B-0126)

0.02g          xylene cyanole FF (Kodak T-1579)

dH<sub>2</sub>O to 10ml (store at -20°C).

**Alkaline lysis solution (NaOH/SDS)**

20ml           1 N NaOH

10ml           10% SDS

dH<sub>2</sub>O to 100ml (make fresh)

**9.5M NH<sub>4</sub>OAc (ammonium acetate)**

73.23g          NH<sub>4</sub>OAc

dH<sub>2</sub>O to 100ml

**8.0M NH<sub>4</sub>OAc:**

61.69g          NH<sub>4</sub>OAc

dH<sub>2</sub>O to 100ml

**15% Ammonium persulfate (APS)**

1.5g            APS

dH<sub>2</sub>O to 10ml (store at 4°C).

**Ampicillin (Amp)**

0.5g            Ampicillin (Sigma A-9518)

dH<sub>2</sub>O to 100ml (Add to media for final conc. 100 µg/ml)

**BCA reagent A**

Sodium carbonate

Sodium bicarbonate

Bichinchoninic acid

Sodium tartrate

0.1M            NaOH

**BCA reagent B**

4% sodium sulphate solution

**1mg/ml BSA (bovine serum albumin)**

5mg            BSA (Sigma A-9647)

dH<sub>2</sub>O to 5ml (aliquot and store at -20°C)

**100 mM calcium chloride (CaCl<sub>2</sub>)**

1.48g            CaCl<sub>2</sub>·2H<sub>2</sub>O  
dH<sub>2</sub>O to 100ml  
autoclave to sterilize (store at 4°C)

**50 mM calcium chloride**

0.74g            CaCl<sub>2</sub>·2H<sub>2</sub>O  
dH<sub>2</sub>O to 100ml  
autoclave to sterilize (store at 4°C)

**Cre reaction buffer**

10mM            Tris-HCl, pH 7.7  
10mM            MgCl<sub>2</sub>  
1mM             DTT  
50mM            NaCl

**10X denaturing buffer**

2ml              1 M Tris-HCl, pH 9.5  
20μl             0.5 M EDTA, pH 8.0  
1ml              100 mM spermidine  
dH<sub>2</sub>O to 10ml (aliquot and store at -20°C)

**Denhardt's Reagent (100x)**

2g                Ficoll  
2g                Polyvinyl pyrrolidone  
2g                Bovine serum albumin (Fraction V)  
make up to 100ml with dH<sub>2</sub>O

**1 M DTT (Dithiothreitol, Cleland's reagent)**

1.54g            DTT (Calbiochem 233155)  
dH<sub>2</sub>O to 10ml (aliquot and store at -20°C).

**DNase-free RNase A**

200mg            RNase A (Sigma R-5500)  
3.3μl             3 M NaOAc, pH 4.5  
dH<sub>2</sub>O to 10ml  
boil for 10 minutes (aliquot and store at -20°C).

**0.5 M EDTA, pH 8.0 (disodium ethylenediamine tetraacetate)**

186.1g           Na<sub>2</sub>EDTA  
Dissolve in approx. 400ml dH<sub>2</sub>O, adjust pH to 8.0 with 10 N NaOH, and adjust to 1 litre final volume with distilled water.

**100 mM EDTA**

20ml             0.5 M EDTA  
80ml             dH<sub>2</sub>O

**5mg/ml ethidium bromide (EtBr)**

500mg            EtBr (Sigma E-8751)  
dH<sub>2</sub>O to 100ml

**20% glucose**

20g                    D-glucose  
dH<sub>2</sub>O to 100ml  
filter sterilize

**1 M HEPES, pH 7.5**

23.83g                HEPES (Sigma H-3375)  
dH<sub>2</sub>O to 100ml  
adjust pH to 7.5 with potassium hydroxide (KOH) (store at 4°C).

**IPTG (isopropyl β-D-thiogalactopyranoside)**

250mg                IPTG (Sigma I-5502)  
dH<sub>2</sub>O to 10ml (aliquot and store at -20°C)

**Kanamycin sulfate (Kan)**

0.5g                   Kanamycin Monosulfate (Sigma K 4000)  
dH<sub>2</sub>O to 100ml (Add to media for final conc. 20 µg/ml)

**1M KCl (potassium chloride)**

7.5g                   KCl  
dH<sub>2</sub>O to 100ml

**Labelling Mix**

7.5µM each of dGTP, dCTP, dTTP

**LB Medium**

10g                    Bacto-Tryptone (Difco 0123-01-1)  
5g                      Bacto-yeast extract (Difco 0127-05-3)  
10g                    NaCl  
dH<sub>2</sub>O to 1 L  
adjust the pH to 7.0 and then autoclave to sterilize

**LB plates**

10g                    Bacto-Tryptone (Difco 0123-01-1)  
5g                      Bacto-yeast extract (Difco 0127-05-3)  
10g                    NaCl  
15g                    Bacto-agar (Difco 0140-01)  
dH<sub>2</sub>O to 1 L  
autoclave to sterilize, cool to 55°C, add antibiotic if desired, and pour into sterile petri dishes (approx. 20ml/plate).

**10x Ligation buffer**

5ml                    1 M Tris-HCl, pH 7.6  
1ml                    1 M MgCl<sub>2</sub>  
1ml                    1 M DTT  
1ml                    100 mM ATP  
2.5mg                BSA  
dH<sub>2</sub>O to 10ml (store in 25ml aliquots at -20°C)

**Lysis/Solution P2**

200mM NaOH

1% SDS

Stored at room temperature

**1 M MgCl<sub>2</sub> (magnesium chloride)**

20.33g MgCl<sub>2</sub>·6H<sub>2</sub>O

dH<sub>2</sub>O to 100ml

**1 M MgSO<sub>4</sub> (magnesium sulfate)**

12.04g MgSO<sub>4</sub>

dH<sub>2</sub>O to 100ml (autoclave)

**Methocarn**

4 volumes methanol

2 volumes chloroform

1 volumes glacial acetic acid

**Microinjection Buffer**

100mM Glutamic acid

1mM DTT

140mM KOH

adjust pH to 7.2 with citric acid

**1 M MOPS**

20.93g MOPS (Sigma M-1254)

Dissolve in 80ml dH<sub>2</sub>O, adjust pH to 7.5 with 1 N NaOH, and bring volume to 100ml.

**10X MOPS buffer**

400ml 1 M MOPS, pH 7.5

170ml 3 M NaCl

100ml 1 M MgCl<sub>2</sub>

330ml dH<sub>2</sub>O

**20 mM dNTP stocks:**

80μl 100 mM dNTP

40μl 50:1 TE buffer

280μl dH<sub>2</sub>O

**Neutralisation/Solution P3**

3.0M potassium acetate, pH 5.5

Stored at room temperature

**PBS**

120mM NaCl

2.7mM KCl

10mM phosphate buffer salts

**5x plasmid reaction buffer**

1.0M	Tris-HCl pH 7.5
100mM	MgCl <sub>2</sub>
250mM	NaCl

**Pre-Hybridisation solution**

6x	SSC
5x	Denhardt's reagent
17mM	SDS

**10X PCR buffer**

5ml	1 M KCl
1ml	1 M Tris-HCl, pH 8.5
150μl	1 M MgCl <sub>2</sub>
dH <sub>2</sub> O to 10ml	

**PCR Deoxynucleotide Preparation**

250μl	100 mM dATP
250μl	100 mM dCTP
250μl	100 mM dGTP
250μl	100 mM dTTP
11.5ml	dH <sub>2</sub> O

Aliquot this into 25 tubes with 500μl in each tube.

**Phenol, TE-saturated**

Add an equal volume of 10 mM Tris-HCl, pH 7.5-8.0, 1 mM Na<sub>2</sub>EDTA to ultrapure phenol, mix well, allow phases to separate, remove and discard upper (aqueous) phase. Repeat until the pH of the aqueous phase is between 7.5-8.0 (store at 4° C).

**Phenol/chloroform/isoamyl alcohol (25:25:1)**

100ml	TE-saturated phenol
100ml	chloroform
4ml	isoamyl alcohol

**Buffer QBT (Equilibration buffer)**

750mM	NaCl
50mM	MOPS, pH 7.0
15%	isopropanol
0.15%	Triton® X-100

**Buffer QC (wash buffer)**

1.0M	NaCl
50mM	MOPS, pH 7.0
15%	isopropanol

**Buffer QF (Elution buffer)**

1.25M	NaCl
50mM	Tris, pH 8.5
15%	isopropanol

**Resuspension/Solution P1**

50mM Tris-HCl, pH 8.0  
10mM EDTA  
100µg/ml RNase A  
Stored at 4°C after addition of RNase A

**RIPA Buffer**

50mM NaCl  
1% NP-40  
12mM deoxycholate  
3mM SDS  
50mM Tris-HCl pH 7.5

**Sequenase® enzyme dilution buffer**

10mM Tris-HCl pH 7.5  
5mM DTT  
0.1mM EDTA  
0.5mg/ml Acetylated BSA

**Sequencing Gel Solution**

250g Urea  
75ml 19:1, 40% acrylamide solution  
50ml 10xTBE  
175ml dH<sub>2</sub>O

For each gel 60ml of the above solution was mixed with 60µl of TEMED and 60µl of fresh 25% ammonium persulphate (in dH<sub>2</sub>O) and immediately poured.

**Sequencing Stop Solution**

95% v/v formamide  
20mM EDTA  
0.5g/l bromophenol blue  
0.5g/l xylene cyanol FF

**Sequencing Termination Mixes**

one of each of:

80µM dATP + 8µM ddATP  
80µM dCTP + 8µM ddCTP  
80µM dGTP + 8µM ddGTP  
80µM dTTP + 8µM ddTTP

**SOC Media**

2% w/v Tryptone  
0.5% w/v Yeast Extract  
0.05% NaCl

Autoclave and add the following filter sterilised reagents. 1% (v/v) 1M MgCl<sub>2</sub>, 1% (v/v) 1M MgSO<sub>4</sub> and 0.1% (v/v) 2M glucose solution.



**2M NaOAc (sodium acetate)**27.22g NaOAc-3H<sub>2</sub>OdH<sub>2</sub>O to 100ml**3M NaOAc, pH 4.5**408.24g NaOAc-3H<sub>2</sub>O

Dissolve in approx. 800ml dH<sub>2</sub>O, adjust pH to 4.5 with glacial acetic acid and bring to a final volume of 1 L with dH<sub>2</sub>O.

**3M NaCl (sodium chloride)**

17.53g NaCl

dH<sub>2</sub>O to 100ml**10N NaOH (sodium hydroxide)**

40g NaOH

dH<sub>2</sub>O to 100ml.**1N NaOH**

10ml 10 N NaOH

dH<sub>2</sub>O to 100ml**10% SDS (sodium dodecyl sulfate)**

10g SDS (Fisher S529-3)

dH<sub>2</sub>O to 100ml**20X SSC (standard saline-citrate:**

17.53g NaCl

8.82g sodium citrate

Dissolve in approx. 80ml dH<sub>2</sub>O, adjust pH to 7.0 with HCl and bring final volume to 100ml.

**1X SSC (standard saline-citrate)**

5ml 20X SSC

95ml dH<sub>2</sub>O**T4 DNA polymerase reaction buffer**

50mM NaCl

10mM Tris-HCl

10mM MgCl<sub>2</sub>

1mM DTT

pH 7.9 at 25°C

**20X TAE buffer**

96.9g Tris base

32.8g NaOAc-3H<sub>2</sub>O14.9g Na<sub>2</sub>EDTA

Dissolve in approx. 700ml of deionised water, adjust the pH to 8.3 with glacial acetic acid, and bring to 1 L with dH<sub>2</sub>O.

**10X TBE**

216g            Tris base  
110g            boric acid  
16.6g           EDTA  
Add water to 2 litres.

**5x TBS**

200g            NaCl  
5g               KCl  
75g              Tris-HCl  
make up to 4L with dH<sub>2</sub>O, pH to 7.4 with HCl  
make up to 5L

**5x TBST**

200g            NaCl  
5g               KCl  
75g              Tris-HCl  
make up to 4L with dH<sub>2</sub>O, pH to 7.4 with HCl  
add 12.5ml    Tween-20  
make up to 5L

**TEMED (N,N,N',N'-tetramethylethylenediamine)**

Severn 20-3000-01, store protected from light at 15°C.

**TE (10:1) buffer**

10ml            1 M Tris-HCl, pH 7.6  
2ml              0.5 M EDTA  
dH<sub>2</sub>O to 1 L

**Tetracycline stock (Tet)**

1g                Tetracycline (Sigma T-3383)  
50ml            100% ethanol  
dH<sub>2</sub>O to 100ml (store at 4°C in the dark)  
Add to media for final conc. 20µg/ml.

**1M Tris-HCl, pH 7.6, 8.0, 8.5, 9.0, 9.5**

121.1g           Tris base  
dH<sub>2</sub>O to 800ml  
Adjust pH with concentrated HCl and then add dH<sub>2</sub>O to 1 L.

**X-gal (5-bromo-4-chloro-3-indolyl b-D-galactopyranoside):**

200mg           X-gal (Sigma B-4252)  
dimethylformamide (DMF) to 10ml  
Aliquot and store protected from light at -20°C

## **Appendix 2**

### ***Cdx-1* promoter sequence data**

Positions are listed relative to the transcription start site

-6456	CTCGAGAGAT	TGGGGGGGGG	GGCTGCCATT	TCCGGCCAGC	CCAGCCATCC
	GAGCTCTCTA	ACCCCCCCCC	CCGACGGTAA	AGGCCGGTCG	GGTCGGTAGG
-6406	GCTGGCCCAT	CTGGAAGCAC	TTAGTTCTAC	CACCATTAAAC	CCTTGGTGGG
	CGACCGGGTA	GACCTTCGTG	AATCAAGATG	GTGGTAATTG	GGAACCACCC
-6356	CACAGGAGCT	GTTTCTCCTC	CCTGTGATGG	GATGCAAGCG	ATACAAAACCT
	GTGTCCCTCGA	CAAAGAGGAG	GGACACTACC	CTACGTTCGC	TATGTTTTGA
-6306	ACACAGAGTG	CCACAGGGCA	CTCTCGAAGG	CCTGGACCCA	GAGACTAAAG
	TGTGTCTCAC	GGTGTCCCGT	GAGAGCTTCC	GGACCTGGGT	CTCTGATTTCT
-6256	TTGAGTTCTG	ACTCCAGCCT	CTCTGTTTGA	AGCTGGAAGA	CCTTGTGTGAA
	AACTCAAGAC	TGAGGTCGGA	GAGACAAACT	TCGACCTTCT	GGAACAACCTT
-6206	TGTCCCTGAG	CTGCCCAGCC	TGCTTTCTGC	TCTAGACAGC	AGGTAAGTCA
	ACAGGGACTC	GACGGGTCGG	ACGAAAGACG	AGATCTGTCTG	TCCATTTCAGT
-6156	GGGCCGGGTC	TGCACCCGCA	GCTTCTTCAG	CAGTCTGTGG	GTGAACGAGT
	CCCGGCCAG	ACGTGGGCGT	CGAAGAAGTC	GTCAGACACC	CACCTTGCTCA
-6106	ATGTAGCATC	ACGCTAGGGT	GGCCAGCATG	GTCCCAGCCC	CACCTGTGGAC
	TACATCGTAG	TGCGATCCCA	CCGGTTCGTAC	CAGGGTCGGG	GTGACACCTG
-6056	AGAACGAATC	AAGCACACTC	ATTCCACGCC	CAGATCCTGG	GAGCCGGTTT
	TCTTGCTTAG	TTCGTGTGAG	TAAGGTGCGG	GTCTAGGACC	CTCGGCCAAA
-6006	ACCTATGGTC	AAACACTGAG	GGAAAAGCTG	CTGAGAACGG	GTGGTGGGTA
	TGGATACCAG	TTTGTGACTC	CCTTTTCGAC	GACTCTTGCC	CACCACCCAT
-5956	GAGTTCAAAA	GGGGCTCCAG	CCTTGGGGCG	TCACGCTTTT	TTCTACCTAC
	CTCAAGTTTT	CCCCGAGGTC	GGAACCCCGC	AGTGCGAAAA	AAGATGGATG
-5906	CCAGAATGGC	GGTCTACTTC	ATGCATGAGA	ATCTGAGCGC	TAAGCCTGAG
	GGTCTTACCG	CCAGATGAAG	TACGTACTCT	TAGACTCGCG	ATTCCGACTC
-5856	GCTCCGAACCT	CAGTTGCGTC	TTAGCTGTGT	GACTCTCAAA	ATGCCCAGAG
	CGAGGCTTGA	GTCAACGCAG	AATCGACACA	CTGAGAGTTT	TACGGGTCTC
-5806	CCTTGGGGTT	TTTAAAATAA	AGGCTTGTGT	GCCCACCTCA	CATGCTGTGA
	GGAACCCCAA	AAATTTTATT	TCCGAACACA	CGGGTGGAGT	GTACGACACT
-5756	GGGACATGAG	TACAAGATTTC	ATACAGGTTT	GGGGCGAGCC	TAGAGGCAGA
	CCCTGTACTC	ATGTTCTAAG	TATGTCCAAA	CCCCGCTCGG	ATCTCCGTCT
-5706	AGGAGAGGAG	AACTACAAAG	GATAACCAGT	CCTCCCTGGG	GCTGCTGTCA
	TCCTCTCCTC	TTGATGTTTC	CTATTGGTCA	GGAGGGACCC	CGACGACAGT
-5656	CAGGAGTCCA	TAATCAGGGC	CGGAAGGACC	CCCCTCCACC	CACCTCTCTTT
	GTCTCAGGT	ATTAGTCCCG	GCCTTCCTGG	GGGGAGGTGG	GTGAGAGAAA

-5606	ATAGCCAACG	TAAGGGGTCT	TGACAGAGTG	GCTCATGTGA	GGCTGTATGA
	TATCGGTTGC	ATTCCCCAGA	ACTGTCTCAC	CGAGTACACT	CCGACATACT
-5556	TCCTGCAGTG	TGAGCACAGA	TACCATGGCA	ACAAAGGCTT	GGCCCCAAGT
	AGGACGTCAC	ACTCGTGTCT	ATGGTACCGT	TGTTTCCGAA	CCGGGGTTCA
-5506	TTTATACCCA	CGTTCTGCCC	TAGAGGAGGG	GTCCAGGTTG	CGCATAATCC
	AAATATGGGT	GCAAGACGGG	ATCTCCTCCC	CAGGTCCAAC	GCGTATTAGG
-5456	CTGATACCTC	CCTCACTGTC	AGAGCAGGAA	GCAGGGCTCT	GTGGAGTGTG
	GACTATGGAG	GGAGTGACAG	TCTCGTCCTT	CGTCCCGAGA	CACCTCACAC
-5406	GGCTGAGGGG	CCCGGAAGGG	GATCCCCATG	GGAGATGTCC	CTGTGGATCA
	CCGACTCCCC	GGGCCTTCCC	CTAGGGGTAC	CCTCTACAGG	GACACCTAGT
-5356	GCTCAGCACC	AGCAGCTACC	ACAGCTGGCT	CTCCCACGCT	CCTGAAGCTG
	CGAGTCGTGG	TCGTGATGG	TGTCGACCGA	GAGGGTGCAG	GGACTTCGAC
-5306	CCCTGGAGGC	TCGAGAGCA	GGCATGGGGT	GTGAGGAAGC	AGCTCACGTT
	GGGACCTCCG	AGCGTCTCGT	CCGTACCCCA	CACTCCTTCG	TCGAGTGCAA
-5256	CCAAACTTAC	ACAATCAGNT	CGGTGTCCNG	AAGGACGTCC	TNTAGGTCCT
	GGTTTGAATG	TGTTAGTCNA	GCCACAGGNC	TTCTTGACAG	ANATCCAGAG
-5206	TACTGACAAC	TTAGGAAATG	GAAGGCTTGC	AGGAGGTCCT	GACGGTAAGA
	ATGACTGTTG	AATCCTTTAC	CTTCCGAACG	TCCTCCAGGA	CTGCCATTCT
-5156	TATTAAGCCA	GATCTTCCCT	ATTACTCACT	GTGAGAGATA	GNTAGNCCTN
	ATAATTCGGT	CTAGAAGGGA	TAATGAGTGA	CACTCTCTAT	CNATCNGGAN
-5106	CTANAGCTGT	CTACNGTAGC	TGGAAAGTCA	AGTCTTAGGG	TAGGAGTCAC
	GATNTCGACA	GATGNCATCG	ACCTTTTCAGT	TCAGAATCCC	ATCCTCAGTG
-5056	AGCACCGTTG	ACACTTGGAG	CAGGTGTCCG	AGGAGGCAGC	AGGGACATGA
	TCGTGGCAAC	TGTGAACCTC	GTCCACAGGC	TCCTCCGTCG	TCCCTGTACT
			SmaI		
			~~~~~		
-5006	GTCATCTTCA	ACAGAGATTG	TTGGAAACCC	TCTTCCCGGG	CCAGGCTCCA
	CAGTAGAAGT	TGTCTCTAAG	AACCTTTGGG	AGAAGGGCCC	GGTCCGAGGT
-4956	CTAAAGCAGG	ACATTTGAAG	TTCAAGACAT	GTCCATTTCAT	ACTTACCCCA
	GATTTGCTCC	TGTAAACTTC	AAGTTCTGTA	CAGGTAAGTA	TGAATGGGGT
-4906	GCCTCCCAAC	CACCCTATCC	TATGCTACCT	CAGTGCCCCC	GGAACACAGG
	CGGAGGGTTG	GTGGGATAGG	ATACGATGGA	GTCACGGGGG	CCTTGTGTCC
-4856	GCACTGTGGC	ACATCCTCCT	ACCATATACA	GAAGTCTCAA	AGAAATACAG
	CGTGACACCG	TGTAGGAGGA	TGGTATATGT	CTTCAGAGTT	TCTTTATGTC

-4806 CAGTAGCACC CTTGAAGACG ATATTGTTTC TCCCTCACGA GAACACTGGC  
 GTCATCGTGG GAACTTCTGC TATAACAAAG AGGGAGTGCT CTTGTGACCG  
  
 -4756 TCTTAAGACC CAGAAGTGGG TACAGAACCT GACTGAAGGC ACATAATATG  
 AGAATTCTGG GTCTTCACCC ATGTCTTGGA CTGACTTCCG TGTATTATAC  
  
 -4706 TCTTCCGTGG GGCTCTCAC CCCTAATTTT AGGCTGTCCT GTTTCTCCCA  
 AGAAGGCACC CCGGAGAGTG GGGATTAAAG TCCGACAGGA CAAAGAGGGT  
  
 -4656 TTGTGCATCT GCATGAAGGG AGTTAGTTTC CTGAACACCT ACTATGTGAC  
 AACACGTAGA CGTACTTCCC TCAATCAAGC GACTTGTGGA TGATACACTG  
  
 -4606 CTGCCTCCCC ACGCCCTTTC CCAGAGCTCA CAAGCAAACA TCAAGGAAGC  
 GACGGAGGGG TCGGGGAAG GGTCTCGAGT GTTCGTTTGT AGTTCCTTCG  
  
 -4556 AACGGATGAA TAACACCAAT CTGTGGTTGG TACAATGAAA GATACAGAAA  
 TTGCCTACTT ATTGTGGTTA GACACCAACC ATGTTACTTT CTATGTCTTT  
  
 -4506 TCACAGAGTT GGGCAGATGG GGGTGGGGGG TCATGCTCAG GCCTGTAGAA  
 AGTGTCTCAA CCCGTCTACC CCCACCCCCC AGTACGAGTC CGGACATCTT  
  
 -4456 GCTCCAAGGG NATAACGCTC TTTGGGGAAG ACGTATCTCC GCTCAGACCT  
 CGAGGTTCCC NTATTGCGAG AAACCCCTTC TGCATAGAGG CGAGTCTGGA  
  
 -4406 GAGGTTGAAA GGCCATCCAT GAAACAAGAA TGGGAACAGT GTTGAAAGCA  
 CTCCAACTTT CCGGTAGGTA CTTTGTTCCTT ACCCTTGTC CAACTTTCGT  
  
 -4356 AAGGGATCAG CCCCTGCAAA TTCTTTGAGG TTCACAAGAC CTTGAAGTGG  
 TTCCCTAGTC GGGGACGTTT AAGAACTCC AAGTGTTCTG GAACTTCACC  
  
 -4306 TAGATCAGAA GAGAGGAGAG GGATTGCCAG GATGGGGAAG GCAGCTGGGG  
 ATCTAGTCTT CTCTCCTCTC CCTAACGGTC CTACCCCTTC CGTCGACCCC  
  
 -4256 GCCAGCGCAA AGACTTGAGG GCTCTTGGGA GCCCAGCTTT TATTCTGAGG  
 CGGTGCGGTT TCTGAACTCC CGAGAACCCT CGGGTCGAAA ATAAGACTCC

SphI

~~~~~

-4206 TCAGCAAGCC GCTGCAGCTC AGCATGCATG AGACCAGGTA GCTGAGGTAT  
 AGTCGTTCCG CGACGTCGAG TCGTACGTAC TCTGGTCCAT CGACTCCATA  
  
 -4156 AGTGTGGGGG TGTAAGGCTG GGCTGGAGTC TTGGGGAGCT TGTGGGCAAC  
 TCACACCCCC ACATTCCGAC CCGACCTCAG AACCCTCGA ACACCCGTTG  
  
 -4106 TGTGAATCAG AGTTGGAACC TCTAGGCTCT CTCTGAGCTT TGACCCCTGA  
 ACACTTAGTC TCAACCTTGG AGATCCGAGA GAGACTCGAA ACTGGGGACT  
  
 -4056 GTCTCCGTGC CTCCTTTCCT CCCCTGGCCT GTCTGTTTGC ACATTGTTGG  
 CAGAGGCACG GAGGAAAGGA GGGGACCGGA CAGACAAACG TGTAACAACC



~~~~~  
 -4006 GCACAAGGAC ATTCTGTGCT TGTCTTGAGC CCTGATTCTG GTGTCTTCTT  
 CGTGTTCCCTG TAAGACACGA ACAAGACTCG GGACTAAGAC CACAGAAGAA

TCF/LEF binding motif

~~~~~

-3956 TGATCTCTGG TGTCTGGATA AGGAGGTTGT CTCAATTCCC CAATACCTCC  
 ACTAGAGACC ACAGACCTAT TCCTCCAACA GAGTTAAGGG GTTATGGAGG

-3906 TGTTATCTGA GACTAGGCAA TAAGGGACTA ATTTCTCCTG TCATGCTCTA  
 ACAATAGACT CTGATCCGTT ATTCCCTGAT TAAAGAGGAC AGTACGAGAT

-3856 AAGGCCAGCA GGTGGCGACA TCCCTCTAAA GGGCTGGAGA ATGACAGCCT  
 TTCCGGTCGT CCACCGCTGT AGGGAGATTT CCCGACCTCT TACTGTCGGA

SmaI

~~~~~

-3806 GTCTATAAAC CAAAATTTTC CCGGGACCCT ACTGTCGATA CNNGCGATNG  
 CAGATATTTG GTTTTAAAG GGCCCTGGGA TGACAGCTAT GNNCGCTANC

-3756 CGNGGCTCAT NTTTGCCTCT GNGANGCTGA CAGCCNTNCN TGTTTTNNGA  
 GCNCCGAGTA NAAACGGAGA CNCTNCGACT GTCGGNANGN ACAAANNCT

-3706 GGAGAACCGG TGACCCCNAG AAGGGAACTA NCCTNCCNAG GGNCAGANAG  
 CCTCTTGGCC ACTGGGGNTC TTCCCTTGAT NGGANGGNTC CCNGTCTNTC

-3656 CTTGGAAGAA NCTNGTNTGA CCCTTGGTAA NATNTCNGNA NAGGGACTCN  
 GAACCTTCTT NGANCANACT GGAACCAT NTANAGNCNT NTCCCTGAGN

-3606 ANTATGNTAG NNGNGGGGGC ATGGTGGGTN TGACACTANT CCCTGGTGTN  
 TNATACNATC NNCNCCCCCG TACCACCCAN ACTGTGATNA GGGACCACAN

-3556 NANTGGCTTT CNTGAAGCAA ACCTAATNTN AGCAGGAACT GANCCCCAGA  
 NTNACCGAAA GNACTTCGTT TGGATTANAN TCGTCCTTGA CTNNGGGTCT

-3506 CCTTNCNCTT NCAGGGCCAT AAACCAAGTT ATCANNACNA TCTTTATANG  
 GGAANGNGAA NGTCCCGGTA TTTGGTTCAA TAGTNNTGNT AGAAATATNC

-3456 GCCTNNATNT GGGTACCCCA CCCCAGGAGG CCCCAGGAGTG AGAGCCTCCA  
 CGGANNTANA CCCATGGGGT GGGGTCTCTC GGGGCCTCAC TCTCGGAGGT

-3406 GGGGGAAGAG GGAAGAAGGA GGGGTCCAAT AAACAGCCAC CCCCACAGAT  
 CCCCCTTCTC CTTCTTCTCT CCCCAGGTTA TTTGTCTGGT GGGGTGTCTA

-3356 AACGGGATTG CTTGTAAACA GCCACGCCCA GAGATAAGGG GATTGCCTGA  
 TTGCCCTAAC GGACATTTGT CGGTGCGGGT CTCTATTCCC CTAACGGACT

-3306 AAGGAAGAGG CTGCTTCAGG ATAGGCGGGG CCGGCACAGT GGTACAGTTA  
 TTCCTTCTCC GACGAAGTCC TATCCGCCCC GGCCGTGTCA CCATGTCAAT

-3256 ACAGTCCCGG AGGGCACATG TGAAACTGGT GTCCAGGGGC TCAAGATCGG  
 TGTCAGGGCC TCCCGTGTAC ACTTTGACCA CAGGTCCCCG AGTTCTAGCC

-3206 GTCGCAGGTC AGAGGTGGGG GTGTACGGCA CTACAATCTC AGACTTAAAT  
CAGCGTCCAG TCTCCACCCC CACATGCCGT GATGTTAGAG TCTGAATTTA

-3156 AAGGTTATAA TTCCTTTCTT TAAGGAACTG ATGCTAAAGC CTCTAACCAG  
TTCCAATATT AAGGAAAGAA ATTCCTTGAC TACGATTTTCG GAGATTGGTC

-3106 GGGGTCCACC TAGGTGTGAG GGTGGCTGGA GCCTTGAGGC TGAGGAGGAA  
CCCCAGGTGG ATCCACACTC CCACCGACCT CGGAACTCCG ACTCCTCCTT

-3056 GAGGAACCAG CAATTTGTTA AGCAAAATCC AGCTAACAAG TTGTCAGGTG  
CTCCTTGGTC GTTAAACAAT TCGTTTTAGG TCGATTGTTT AACAGTCCAC

-3006 GGGGTGTGAG TGTGAGGATC ACTCTCTGTG CTCAGTGGGT CAACCATCTC  
CCCCACAGTC ACAGTCCTAG TGAGAGACAC GAGTCACCCA GTTGGTAGAG

-2956 TAGAGGGAAG AAAAAGTACC AGCTAGGTGG GTTTTCTGTG TGAAGTCTAGG  
ATCTCCCTTC TTTTGTGACG TCGATCCACC CAAAAGACAC ACTGAGATCC

-2906 AAGGTGACTT TACCTCTCTG AACACATTTG TTAATCTGGG GAAAAAATG  
TTCCACTGAA ATGGAGAGAC TTGTGTAAAC AATTAGACCC CTTTTTTTAC

-2856 GGGGTAGCAA TCGCTTTTTT TCCCTTGATG GTGGTCANGA TGAACCGAGG  
CCCCATCGTT AGCGAAAAAA AGGGAAGTAC CACCAGTNCT ACTTGGCTCC

-2806 ACACACGCTG GCCGTGAGAT GTCACCTCTCA GCACCTGCTC AGTGTGTGGA  
TGTGTGCGAC CGGCACCTCA CAGTGAGAGT CGTGGACGAG TCACACACCT

-2756 AGTGCCATTG TCACAAACCA GTGTGCCATC TCAGAGAGAG ACAGCGGATA  
TCACGGTAAC AGTGTTTGGT CACACGGTAG AGTCTCTCTC TGTCGCCAT

-2706 CTCAGAAAGG GTGGTGCAGC CCAGGAGCAC CTGTAGATGG GGGGTGGGGG  
GAGTCTTTCC CACCACGTCG GGTCTCTGTG GACATCTACC CCCCACCCCC

-2656 GCGGTATGTC CACCCTTTGA GGGTACAGCA TTGAGAAACG TCTCTGCAGC  
CGCCATACAG GTGGGAAACT CCCATGTCGT AACTCTTTGC AGAGACGTCG

-2606 AGGCTTCTGA ATTTCTCTATA AGGTCCCCTG CCACCCTGAC CCATCCATCT  
TCCGAAGACT TAAAGGATAT TCCAGGGGAC GGTGGGACTG GGTAGGTAGA

-2556 CTAAGCCACT ACCAAATCCT GGCTATGTCT CAGTCACTTA GGATCCCTGC  
GATTCGGTGA TGGTTTAGGA CCGATACAGA GTCAGTGAAT CCTAGGGACG

-2506 CCAGCTTCCC TCTCCTGGTG CAGGGGTTAG GAGGGGTGGG TGTGTTCTCA  
GGTCGAAGGG AGAGGACCAC GTCCCCAATC CTCCCCACCC ACACAAGAGT

-2456 GATCAGCTTC CTGCTTCTCC CTCAGGGCTG GGTATCAGT TCTAGGGGAA  
CTAGTCGAAG GACGAAGAGG GAGTCCCGAC CCAATAGTCA AGATCCCCCT

-2406 CACACATTCC TCGCCCCCTA AGGCCTTGGC GCACACTGAC CACTGGACTA  
GTGTGTAAGG AGCGGGGGAT TCCGGAACCG CGTGTGACTG GTGACCTGAT

-2356 CACTGCCCTT GTCTACTCGC CCCGACTGAT GTGTCACCTC CTTCCTAAAA  
GTGACGGGAA CAGATGAGCG GGGCTGACTA CACAGTGGAG GAAGGGTTTT

EcoRI  
~~~~~

-2306 GAATTCTCTA GCTTGGCTCC CAACTAGATC CAAGATGTGA TATTCCTTC  
CTTAAGAGAT CGAACCGAGG GTTGATCTAG GTTCTACACT ATAAGGGAAG

-2256 ACTGAACATT TAATTCACGC CTTGCCCCC AGGACTGAAG GCACCTTAAG  
TGAATTGTAA ATTAAGTGCG GAACGGGGG TCCTGACTTC CGTGGAATTC

-2206 TTAGAAGGGG CTTTATTCCA TATTGAACAT GGCGCCAGCA CGGTTGTGTG  
AATCTTCCCC GAAGTAAGGT ATAACCTTGT CCGCGGTCGT GCCAACACAC

-2156 GGTCAGTCTG TCCGTGTCAA AGGGATGAAT ATGTGAGCAA ATAAATTTCT  
CCAGTCAGAC AGGCACAGTT TCCCTACTTA TACACTCGTT TATTTAAAGA

-2106 AATCCCAAAT GGCAAGCCTC TTGACTAGGC AAGAACTGGC ATAGAGAGGA  
TTAGGGTTTA CCGTTCGGAG AACTGATCCG TTCTTGACCG TATCTCTCCT

-2056 TTAGAGGCCA CCAATTTCTT GCTTTCCCCA GAAAAAGTGT GACCCCTCAC  
AATCTCCGGT GGTAAAGGA CGAAAGGGGT CTTTTTCACA CTGGGGAGTG

-2006 TGGAGATACC CAGCCATCCC TTGCACACCT CCCTTGACGA TAAACTCACT  
ACCTCTATGG GTCGGTAGGG AACGTGTGGA GGGAACTGCT ATTTGAGTGA

TCF/LEF binding motif

~~~~~

-1956 CACTACAGGA CCATCCGTCT GCTTTGTACT GGATGGTCCT ACAGTGAGCA  
GTGATGTCCT GGTAGGCAGA CGAAACATGA CCTACCAGGA TGTCACCTCGT

-1906 TTCATTCTAG ACAGGTAGGT TAGACTCCAT GCCTGCAGCT CTCTCTCTCA  
AAGTAAGATC TGTCCATCCA ATCTGAGGTA CGGACGTCGA GAGAGAGAGT

-1856 GGTCCCTGCC CTCCATGCTA GGCCTTCACT TCCCACAGTG AGGAGCAGGC  
CCAGGGACGG GAGGTACGAT CCGGAAGTGA AGGGTGTCAC TCCTCGTCCG

EcoRI

~~~~~

-1806 CTGGTTACCC CAGGTGAGAG CAGAATTGAA TTCTGCCCTC TTGGGGTATC  
GACCAATGGG GTCCACTCTC GTCTTAACTT AAGACGGGAG AACCCCATAG

-1756 TCTTGGTTGT GCTCTCACTG GAACCCAGAA CTCTAGGTGT CTTCTGACCT  
AGAACCAACA CGAGAGTGAC CTTGGGTCTT GAGATCCACA GAAGACTGGA

-1706 GAGAATGTGG CGCCCTCCC CCATGCTGTG ACACCCATGT GCTCATACCG  
CTCTTACACC GCGGGGAGGG GGTACGACAC TGTGGGTACA CGAGTATGGC

-1656 TCATCTGTGT TGGCGACAAG TAGCTACCTT CCCCTCTCTT GTAAGCAGCC  
AGTAGACACA ACCGCTGTTC ATCGATGGAA GGGGAGAGAA CATTCGTCCG

-1606 ATCATTATAA AGCATGAACT GGAGTTCTGC CCTTTCCTCT CTGGGCCTCC  
TAGTAATATT TCGTACTTGA CCTCAAGACG GGAAAGGAGA GACCCGGAGG

-1556 CTGCTTCCAT CCATTGAGCA GGAAGGGGAC CATTAGGCTT ACTGCTATGT  
GACGAAGGTA GGTAACCTCGT CCTTCCCCTG GTAATCCGAA TGACGATACA

-1506 GTAAGTGTCA CTGATTGAGG ACACCTTAC GGGCAGGTCA TCTGTGCACT  
CATTCACAGT GACTAACTCC TGTGGAAGTG CCCGTCCAGT AGACACGTGA

-1456 CGCACAGGGC TCAGCTGAAT CCTGCCACTA TGAAATTCTA GGTCATTTTG  
GCGTGTCCCG AGTCGACTTA GGACGGTGAT ACTTTAAGAT CCAGTAAAC

-1406 GAACAACCTG GCATTTGCAT TTTGCACTGC ATCTGCCGGC AGGGTACAGC  
CTTGTTGGAC CGTAAACGTA AAACGTGACG TAGACGGCCG TCCCATGTGC

-1356 ATGGCCTCGA CTCCCCTGCT TGCCCCTGTA AAGAGAAAGA AAGTCTAATA  
TACCGGAGCT GAGGGGACGA ACGGGGACAT TTCTCTTTCT TTCAGATTAT

-1306 GCCAAGACCT GTGTCTTGCT TTGAGGCATC GGAAATTCTG CATGGCCCTA  
CGGTTCTGGA CACAGAACGA AACTCCGTAG CTTTAAAGAC GTACCGGGAT

-1256 GAGCACCAGA AATCCATCTC AGAGGACTCC TCTGCACACT TTCCTCCTGG  
CTCGTGGTCT TTAGGTAGAG TCTCCTGAGG AGACGTGTGA AAGGAGGACC

-1206 CCCATAGTCT TCCTTCTGAA TCCGTCCTCA GCAATGCCTT CTCTTCCCAG  
GGGTATCAGA AGGAAGACTT AGGCAGGAGT CGTTACGGAA GAGAAGGGTC

-1156 TCTCAGCTAT GCCTTTCCCT GCTCAGGTAG CTGTGGAATG CCTGCCTGCC  
AGAGTCGATA CGGAAAGGGA CGAGTCCATC GACACCTTAC GGACGGACGG

-1106 TGTCTTCAAC GGTGCAGGAA AGACCCACAT GGACATTTAA AATACCAGCG  
ACAGAAGTTG CCACGTCCTT TCTGGGTGTA CCTGTAAATT TTATGGTCGC

-1056 GGATCCGTTT TCCATTTCCA TTACCAACGA GGATCATCAA GCCAACTGCA  
CCTAGGCAAA AGGTAAAGGT AATGGTTGCT CCTAGTAGTT CGGTTGACGT

-1006 AAATCGCGTT GTTTCTTTTT TTTCCCCCA ATCAGTCAGA GTACACAGAA  
TTTAGCGCAA CAAAGAAAAA AAAGGGGGGT TAGTCAGTCT CATGTGTCTT

#### TCF/LEF binding motif

~~~~~

-956 TATAAACTTT GTTCGATAAT AAATACTCAG GAATTTGTCC CTGTATTTTA  
ATATTTGAAA CAAGCTATTA TTTATGAGTC CTTAAACAGG GACATAAAAT

-906 AAGACTCGGG CTGATGTGTG TAGGGGTAAA ACGACATGAA GCCTGGAATT  
TTCTGAGCCC GACTACACAC ATCCCCATTT TGCTGTACTT CGGACCTTAA

-856 TGTTTTTAAA TTGTTCCGAA AATGCTTTAG GAAAAAAAC AAAAGGCAGT  
ACAAAAATTT AACAAAGGCTT TTACGAAATC CTTTTTTTTG TTTTCCGTCA

-806 AGATTTAAAA AAAAAAAAAA AAAAAAAAAA AAAAGAGGTG ACAGAAACCC  
TCTAAATTTT TTTTTTTTTT TTTTTTTTTT TTTTCTCCAC TGTCTTTGGG

-756 GATAAGTAAT GGAAGCTGCA AGGTAGGTAC ACAATGCAAC TCGGTGTATA  
CTATTCATTA CCTTCGACGT TCCATCCATG TGTTACGTTG AGCCACATAT

-706 TGTATGTTTG AAACATTTCA GCAATGATAA TGTCAAGGTA ATACGCATCA  
ACATACAAAC TTTGTAAAGT CGTTACTATT ACAGTTCCAT TATGCGTAGT

-656 AAAGTTTGTG AAGTTTCCCA AGGTGTCCCC AGAAACCCCC CTGGTCCACA  
TTTCAAACAC TTCAAAGGGT TCCACAGGGG TCTTTGGGGG GACCAGGTGT

-606 CCTTTGAGAA TTTCTCTGGG GAGCCAATAA GAGAAACAGT AGCTTGCAGA  
GGAAACTCTT AAAGAGACCC CTCGGTTATT CTCTTTGTCA TCGAACGTCT

#### Retinoic Acid Response Element

~~~~~

-556 CTGGCTATCT GCATATGCAC ATCACCATAG CTTTCATCTG GGACCCGAAG  
GACCGATAGA CGTATACGTG TAGTGGTATC GAAAGTAGAC CCTGGGCTTC

```

-506  GGTCGTGACC CCTAACCCCA CCACCACCCC TCCCAGAGAC GCCCGCTACC
      CCAGCACTGG GGATTGGGGT GGTGGTGGGG AGGGTCTCTG CGGGCGATGG

-456  TTACAGACGG AACCCCGTT TGAAGTCAGC CTTGCTCTCC GCCTCACCTC
      AATGTCTGCC TTGGGGGCAA ACTTCAGTCG GAACGAGAGG CGGAGTGGAG

-406  AAGCTGGTGG TCACTGCTGA CAGTGTCCCC ATGCTCTGTG GCGGGTGTGA
      TTCGACCACC AGTGACGACT GTCACAGGGG TACGAGACAC CGCCCACT

-356  GGCTCGCCTA GGGTCATGCC ACCACTCCAC CCCGCTCCTC GGAGCCAGTT
      CCGAGCGGAT CCCAGTACGG TGGTGAGGTG GGGCGAGGAG CCTCGGTCAA

-306  TCTCCACCTG TAACCCAGGG GTGGGTGGTG GGGAGGTCCC TGCACCCCC
      AGAGGTGGAC ATTGGGTCCC CACCCACCAC CCCTCCAGGG ACGCTGGGGG

-256  GAAGGAGGAG TCGGAACCCC AAGCTGGGAC CGGACCGTCA TCCAGGCCGC
      CTTCTCTCTC AGCCTTGGGG TTCGACCTG GCCTGGCAGT AGGTCCGGCG

-206  GCCCTCCGGG TCCCCAGCCC GCTGGCCCAT CCACCTCCCG CTTAGGGCGG
      CGGGAGGCCC AGGGGTCGGG CGACCGGGTA GGTGGAGGGC GAATCCCGCC

                                           TCF/LEF binding motif
                                           ~~~~~~
-156  CAATTTGTCT CCTTTTGAAC CCCCTCGCCC GACGGGCTTC CCCCTTTGAT
      GTTAAACAGA GGAAACTTG GGGGAGCGGG CTGCCCAGG GGGGAACTA

                                           TCF/LEF binding motif
                                           ~~~~~~
~
-106  TCGCGGCCCC GAGGCTTCCC CCCGCTTTGA AATGCAAAGC CGCCCGGCTG
      AGCGCCGGGG CTCCGAAGGG GGGCGAAACT TTACGTTTCG GCGGGCCGAC

                                           Tata box
                                           ~~~~~~
-56   GGGCCGCGGA CGGCCGCGG CTATAAAAGG CCGGGGTGGG GCGGGCGCGG
      CCCGGCGCCT GCCGGGCGCC GATATTTTCC GGCCCCACCC CGCCCGCGCC

NotI(-2)Transcription start site(0)
~~~~~|~~~~~
-6    CGGCGGCCGC GGGGCCAGGT GAGCAGTCGC TGGTCGTCGG GCGGGCCGGC
      GCCGCCGGCG CCCCGGTCCA CTCGTCAGCG ACCAGCAGCC CCGCCGGCCG

                                           SphI(+73)
                                           ~~~~~~
+44   CGGGCGCGGC GGCTCCAGGG CCCAGCATGC GCGGGGGACC CTGCGGTCAC
      GCCCGCGCCG CCGAGGTCCC GGGTCGTACG CGCCCCCTGG GACGCCAGTG

Translation start site (+95)
|
+94   CATGTACGTG GGCTATGTGC TGGACAAGGA CTCCTAGCC.....
      GTACATGCAC CCGATACACG ACCTGTTTCT GAGGATCGG.....

```

# **Appendix 3**

## ***CC* mouse database**



| Number     | Sex | Apc Status | Transgene Status | Mating          | Birth    | Fate                  | Death    | Age (Days) |
|------------|-----|------------|------------------|-----------------|----------|-----------------------|----------|------------|
| Founder CC | F   | WT         | POS (HET)        | Micro-injection |          | Breed                 | 30/11/99 |            |
| CC1        | M   | WT         | NEG              | CC x F1 WT      | 10/03/98 | Cull                  |          |            |
| CC2        | M   | WT         | POS (HET)        | CC x F1 WT      | 10/03/98 | Breed                 |          |            |
| 3          | M   | WT         | NEG              | CC x Min        | 30/03/98 | Cull                  |          |            |
| 4          | M   | WT         | POS (HET)        | CC x Min        | 30/03/98 | Breed                 |          |            |
| 5          | F   | WT         | NEG              | CC x Min        | 30/03/98 | Cull                  |          |            |
| 6          | M   | WT         | POS (HET)        | CC x Min        | 30/03/98 | Breed                 |          |            |
| 7          | F   | WT         | NEG              | CC x Min        | 30/03/98 | Cull                  |          |            |
| 8          | M   | WT         | NEG              | CC x Min        | 30/03/98 | Cull                  |          |            |
| 9          | M   | WT         | NEG              | CC x Min        | 30/03/98 | Cull                  |          |            |
| 10         | F   | HET        | POS (HET)        | CC x Min        | 30/03/98 | Gut tumours           | 26/02/99 | 326        |
| 11         | M   | WT         | POS (HET)        | CC x Min        | 30/03/98 | Breed                 |          |            |
| 12         | M   | WT         | NEG              | CC x Min        | 30/03/98 | Cull                  |          |            |
| 13         | F   | WT         | NEG              | CC2 x Min       | 14/05/98 | Found dead            | 16/06/98 | 32         |
| 14         | M   | WT         | NEG              | CC2 x Min       | 14/05/98 | Cull                  |          |            |
| 15         | F   | WT         | NEG              | CC2 x Min       | 14/05/98 | Cull                  |          |            |
| 16         | M   | WT         | POS (HET)        | CC2 x Min       | 19/05/98 | Cull                  |          |            |
| 17         | M   | HET        | NEG              | CC2 x Min       | 19/05/98 | Cull                  |          |            |
| 18         | M   | HET        | POS (HET)        | CC2 x Min       | 19/05/98 | Gut tumours           | 26/02/99 | 277        |
| 19         | M   | HET        | POS (HET)        | CC2 x Min       | 19/05/98 | Breed                 |          |            |
| 20         | F   | WT         | POS (HET)        | CC2 x Min       | 01/06/98 | Cull for MEFs (x Min) |          |            |
| 21         | M   | WT         | NEG              | CC2 x Min       | 01/06/98 | Gut tumours           | 05/02/99 | 244        |
| 22         | M   | HET        | NEG              | CC2 x Min       | 01/06/98 | Gut tumours           | 14/06/99 | 373        |
| 23         | F   | WT         | NEG              | CC2 x Min       | 01/06/98 | Cull for MEFs (x Min) |          |            |
| 24         | F   | HET        | POS (HET)        | CC2 x Min       | 01/06/98 | Breed                 |          |            |
| 25         | F   | WT         | NEG              | CC2 x Min       | 01/06/98 | Found dead            | 06/07/98 | 35         |
| 26         | M   | NT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 27         | M   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 28         | F   | WT         | NEG              | CC x Min        | 05/06/98 | Found dead            | 06/07/98 | 31         |
| 29         | F   | WT         | NEG              | CC x Min        | 05/06/98 | Found dead            | 30/06/98 | 25         |
| 30         | F   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 31         | F   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 32         | M   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 33         | M   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 34         | F   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 35         | M   | WT         | NEG              | CC x Min        | 05/06/98 | Cull                  |          |            |
| 36         | F   | WT         | NEG              | CC x Min        | 05/06/98 | Found dead            | 06/07/98 | 31         |
| 37         | F   | WT         | NEG              | CC11 x Min      | 06/06/98 | Cull                  |          |            |
| 38         | M   | WT         | NEG              | CC11 x Min      | 06/06/98 | Cull                  |          |            |
| 39         | F   | WT         | NEG              | CC11 x Min      | 06/06/98 | Cull                  |          |            |
| 40         | M   | WT         | NEG              | CC11 x Min      | 06/06/98 | Cull                  |          |            |

|    |   |     |           |            |          |                          |          |     |
|----|---|-----|-----------|------------|----------|--------------------------|----------|-----|
| 41 | F | WT  | NEG       | CC11 x Min | 06/06/98 | Cull                     |          |     |
| 42 | F | WT  | NEG       | CC11 x Min | 06/06/98 | Cull                     |          |     |
| 43 | M | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 44 | M | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 45 | M | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 46 | M | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 47 | F | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 48 | F | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 49 | M | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 50 | M | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 51 | F | WT  | NEG       | CC4 x Min  | 08/06/98 | Cull                     |          |     |
| 52 | M | WT  | POS (HET) | CC4 x Min  | 15/06/98 | Cull                     | 16/03/00 | 631 |
| 53 | M | WT  | NEG       | CC4 x Min  | 15/06/98 | Gut tumours              | 05/02/99 | 230 |
| 54 | M | HET | POS (HET) | CC11 x Min | 12/06/98 | Breed                    |          |     |
| 55 | M | HET | POS (HET) | CC11 x Min | 12/06/98 | Gut tumours              | 02/02/99 | 230 |
| 56 | M | HET | NEG       | CC11 x Min | 12/06/98 | Gut tumours              | 19/01/99 | 217 |
| 57 | F | WT  | NEG       | CC11 x Min | 12/06/98 | Cull for MEFs<br>(x Min) |          |     |
| 58 | F | HET | NEG       | CC11 x Min | 12/06/98 | Found dead               | 09/08/99 | 417 |
| 59 | M | HET | POS (HET) | CC11 x Min | 12/06/98 | Breed                    |          |     |
| 60 | F | HET | POS (HET) | CC11 x Min | 12/06/98 | Breed                    |          |     |
| 61 | F | WT  | NEG       | CC11 x Min | 12/06/98 | Cull for MEFs<br>(x CC2) |          |     |
| 62 | F | HET | POS (HET) | CC11 x Min | 12/06/98 | Breed                    |          |     |
| 63 | F | HET | NEG       | CC2 x Min  | 10/06/98 | Gut tumours              | 10/08/99 | 420 |
| 64 | M | WT  | NEG       | CC2 x Min  | 10/06/98 | Cull                     | 10/09/98 | 90  |
| 65 | M | HET | NEG       | CC2 x Min  | 10/06/98 | Cull                     |          |     |
| 66 | F | WT  | POS (HET) | CC2 x Min  | 10/06/98 | Cull                     |          |     |
| 67 | F | HET | POS (HET) | CC2 x Min  | 10/06/98 | Gut tumours              | 02/02/99 | 232 |
| 68 | F | WT  | POS (HET) | CC2 x Min  | 10/06/98 | Cull                     |          |     |
| 69 | F | WT  | POS (HET) | CC2 x Min  | 10/06/98 | Cull                     |          |     |
| 70 | F | WT  | POS (HET) | CC2 x Min  | 10/06/98 | Cull                     |          |     |
| 71 | F | WT  | POS (HET) | CC2 x Min  | 10/06/98 | Cull                     |          |     |
| 72 | F | WT  | NEG       | CC2 x Min  | 10/06/98 | Cull                     | 10/09/98 | 90  |
| 73 | M | WT  | NEG       | CC2 x Min  | 10/06/98 | Cull                     | 10/09/98 | 90  |
| 74 | M | WT  | NEG       | CC10 x Min | 29/06/98 | Cull                     | 13/03/00 | 614 |
| 75 | F | HET | NEG       | CC10 x Min | 29/06/98 | Found dead               | 03/09/99 | 427 |
| 76 | F | WT  | NEG       | CC10 x Min | 29/06/98 | Cull                     |          |     |
| 77 | M | HET | NEG       | CC10 x Min | 29/06/98 | Found dead               | 03/01/00 | 544 |
| 78 | F | HET | POS (HET) | CC10 x Min | 29/06/98 | Gut tumours              | 02/02/99 | 213 |
| 79 | M | HET | NEG       | CC11 x Min | 03/07/98 | Missing                  |          |     |
| 80 | M | WT  | NEG       | CC11 x Min | 03/07/98 | Gut tumours              | 05/02/99 | 212 |
| 81 | F | HET | NEG       | CC11 x Min | 03/07/98 | Gut tumours              | 26/02/99 | 233 |
| 82 | M | HET | POS (HET) | CC11 x Min | 03/07/98 | Breed                    |          |     |

|     |   |     |           |             |          |                  |          |     |
|-----|---|-----|-----------|-------------|----------|------------------|----------|-----|
| 83  | F | HET | POS (HET) | CC11 x Min  | 03/07/98 | Breed            |          |     |
| 84  | F | HET | NEG       | CC4 x Min   | 09/07/98 | Gut tumours      | 16/11/99 | 487 |
| 85  | F | WT  | NEG       | CC4 x Min   | 09/07/98 | Cull             |          |     |
| 86  | F | HET | NEG       | CC4 x Min   | 09/07/98 | Gut tumours      | 21/06/99 | 342 |
| 87  | M | HET | POS (HET) | CC4 x Min   | 09/07/98 | Breed            |          |     |
| 88  | M | HET | NEG       | CC4 x Min   | 09/07/98 | Cull             |          |     |
| 89  | M | WT  | NEG       | CC4 x Min   | 09/07/98 | Cull             |          |     |
| 90  | M | WT  | POS (HET) | CC4 x Min   | 09/07/98 | Cull             |          |     |
| 91  | F | WT  | POS (HET) | CC4 x Min   | 09/07/98 | Breed            |          |     |
| 92  | M | WT  | POS (HET) | CC11 x Min  | 14/07/98 | Cull             |          |     |
| 93  | F | WT  | NEG       | CC11 x Min  | 14/07/98 | Cull             |          |     |
| 94  | M | HET | NEG       | CC11 x Min  | 14/07/98 | Cull             |          |     |
| 95  | M | HET | NEG       | CC11 x Min  | 14/07/98 | Cull             |          |     |
| 96  | F | WT  | NEG       | CC11 x Min  | 14/07/98 | Cull             |          |     |
| 97  | F | WT  | POS (HET) | CC11 x Min  | 14/07/98 | Cull             | 16/03/00 | 602 |
| 98  | M | HET | POS (HET) | CC11 x Min  | 14/07/98 |                  |          |     |
| 99  | M | WT  | NEG       | CC11 x Min  | 14/07/98 | Cull             | 13/03/00 |     |
| 100 | M | HET | NEG       | CC11 x Min  | 14.07.98 |                  |          |     |
| 101 | M | WT  | NEG       | CC10 x CC18 | 19/08/98 | Cull             | 13/03/00 |     |
| 102 | F | HET | POS (HOM) | CC10 x CC18 | 19/08/98 | Cull for embryos |          |     |
| 103 | M | WT  | NEG       | CC10 x CC18 | 19/08/98 | Cull             | 13/03/00 |     |
| 104 | F | HET | NEG       | CC10 x CC18 | 19/08/98 |                  |          |     |
| 105 | F | HET | POS (HOM) | CC10 x CC18 | 19/08/98 | Cull for embryos |          |     |
| 106 | F | HET | POS (HOM) | CC10 x CC18 | 19/08/98 | Cull embryos     |          |     |
| 107 | M | HET | NEG       | CC55 x CC62 | 26/08/98 | Gut tumours      | 04/02/99 | 158 |
| 108 | M | HET | POS (HOM) | CC55 x CC62 | 26/08/98 | Gut tumours      | 21/01/99 | 145 |
| 109 | M | HET | NEG       | CC55 x CC62 | 26/08/98 | Gut tumours      | 21/01/99 | 145 |
| 110 | F | WT  | POS (HOM) | CC55 x CC62 | 26/08/98 | Cull             |          |     |
| 111 | F | HET | POS (HOM) | CC55 x CC62 | 26/08/98 | Cull for embryos |          |     |
| 112 | M | WT  | NEG       | CC59 x CC67 | 02/09/98 | Cull             | 05/02/99 | 153 |
| 113 | M | WT  | NEG       | CC59 x CC67 | 02/09/98 | Cull             | 05/02/99 | 153 |
| 114 | M | HET | POS (HOM) | CC59 x CC67 | 02/09/98 | Cull             | 16/03/00 |     |
| 115 | M | WT  | POS (HOM) | CC59 x CC67 | 02/09/98 | Found dead       | 05/02/99 | 153 |
| 116 | M | HET | POS (HOM) | CC59 x CC67 | 02/09/98 | Gut tumours      | 07/02/00 | 515 |
| 117 | M | HET | POS (HOM) | CC59 x CC67 | 02/09/98 | Gut tumours      | 14/06/99 | 282 |
| 118 | F | HET | POS (HOM) | CC59 x CC67 | 02/09/98 | Gut tumours      | 22/06/99 | 290 |
| 119 | F | HET | POS (HOM) | CC54 x CC60 | 07/09/98 | Cull for embryos |          |     |
| 120 | M | HET | POS (HOM) | CC54 x CC60 | 07/09/98 | Cull             |          |     |
| 121 | M | HET | POS (HOM) | CC54 x CC60 | 07/09/98 | Found dead       | 30/08/99 | 353 |
| 122 | F | HET | POS (HOM) | CC54 x CC60 | 07/09/98 | Cull             |          |     |
| 123 | F | HET | NEG       | CC54 x CC60 | 07/09/98 | Cull             |          |     |
| 124 | F | HET | POS (HOM) | CC19 x CC24 | 07/09/98 | Cull for embryos |          |     |
| 125 | F | WT  | POS (HOM) | CC19 x CC24 | 07/09/98 | Cull             | 05/02/99 | 148 |

|     |   |     |           |             |          |                  |          |     |
|-----|---|-----|-----------|-------------|----------|------------------|----------|-----|
| 126 | F | WT  | POS (HOM) | CC19 x CC24 | 07/09/98 | Cull             | 05/02/99 | 148 |
| 127 | F | WT  | NEG       | CC19 x CC24 | 07/09/98 | Cull             |          |     |
| 128 | F | WT  | POS (HOM) | CC10 x CC18 | 11/09/98 | Cull             |          |     |
| 129 | M | WT  | POS (HOM) | CC10 x CC18 | 11/09/98 | Cull             |          |     |
| 130 | F | HET | NEG       | CC10 x CC18 | 11/09/98 | Cull             |          |     |
| 131 | M | WT  | POS (HOM) | CC10 x CC18 | 11/09/98 | Found dead       | 05/02/99 | 144 |
| 132 | M | WT  | POS (HOM) | CC10 x CC18 | 11/09/98 | Found dead       | 05/02/99 | 144 |
| 133 | F | HET | POS (HOM) | CC10 x CC18 | 11/09/98 | Cull for embryos |          |     |
| 134 | M | WT  | POS (HOM) | CC10 x CC18 | 11/09/98 | Found dead       | 05/02/99 | 144 |
| 135 | F | HET | POS (HOM) | CC55 x CC62 | 20/09/98 | Cull for embryos |          |     |
| 136 | F | HET | POS (HOM) | CC55 x CC62 | 20/09/98 | Cull             |          |     |
| 137 | M | WT  | POS (HOM) | CC55 x CC62 | 20/09/98 | Found dead       | 16/03/99 | 176 |
| 138 | F | HET | POS (HOM) | CC55 x CC62 | 20/09/98 |                  |          |     |
| 139 | F | HET | POS (HOM) | CC55 x CC62 | 20/09/98 | Cull             |          |     |
| 140 | M | HET | POS (HOM) | CC55 x CC62 | 20/09/98 | Gut tumours      | 01/09/99 | 341 |
| 141 | F | HET | POS (HOM) | CC55 x CC62 | 20/09/98 | Found dead       | 20/10/98 | 30  |
| 142 | F | WT  | POS (HOM) | CC55 x CC62 | 20/09/98 | Cull             |          |     |
| 143 | M | HET | POS (HET) | CC91 x Min  | 23/09/98 | Breed, Cull      | 18/12/98 | 85  |
| 144 | M | WT  | POS (HET) | CC91 x Min  | 23/09/98 | Found dead       | 20/10/98 | 27  |
| 145 | F | WT  | POS (HET) | CC91 x Min  | 23/09/98 | Found dead       | 20/10/98 | 27  |
| 146 | F | WT  | POS (HET) | CC91 x Min  | 23/09/98 | Found dead       | 20/10/98 | 27  |
| 147 | F | WT  | POS (HET) | CC91 x Min  | 23/09/98 |                  |          |     |
| 148 | M | WT  | NEG       | CC91 x Min  | 23/09/98 | Cull             | 05/02/99 | 132 |
| 149 | M | HET | POS (HET) | CC91 x Min  | 23/09/98 | Gut tumours      | 02/09/99 | 339 |
| 150 | M | HET | POS (HOM) | CC54 x CC60 | 01/10/98 | Cull             |          |     |
| 151 | M | WT  | POS (HOM) | CC54 x CC60 | 01/10/98 | Cull             |          |     |
| 152 | M | HET | POS (HOM) | CC54 x CC60 | 01/10/98 | Cull             |          |     |
| 153 | M | HET | POS (HOM) | CC54 x CC60 | 01/10/98 | Cull             |          |     |
| 154 | M | HET | POS (HOM) | CC59 x CC67 | 02/10/98 | Cull             |          |     |
| 155 | F | HET | POS (HOM) | CC59 x CC67 | 02/10/98 | Gut tumours      | 18/06/99 | 256 |
| 156 |   |     |           | No animal   |          |                  |          |     |
| 157 | M | WT  | POS (HOM) | CC59 x CC67 | 02/10/98 | Cull             |          |     |
| 158 | F | HET | NEG       | CC59 x CC67 | 02/10/98 | Cull             |          |     |
| 159 | M | WT  | POS (HOM) | CC59 x CC67 | 02/10/98 | Cull             |          |     |
| 160 | M | HET | POS (HOM) | CC59 x CC67 | 02/10/98 | Cull             |          |     |
| 161 | M | HET | NEG       | CC19 x CC24 | 04/10/98 | Cull             |          |     |
| 162 | F | HET | POS (HOM) | CC19 x CC24 | 04/10/98 | Cull             |          |     |
| 163 | M | WT  | POS (HOM) | CC19 x CC24 | 04/10/98 | Cull             |          |     |
| 164 | M | WT  | POS (HOM) | CC10 x CC18 | 05/10/98 | Found dead       | 09/11/98 | 34  |
| 165 | F | HET | POS (HOM) | CC10 x CC18 | 05/10/98 | Cull             |          |     |
| 166 | M | WT  | POS (HOM) | CC10 x CC18 | 05/10/98 | Cull             |          |     |
| 167 | M | HET | POS (HOM) | CC10 x CC18 | 05/10/98 | Cull             |          |     |
| 168 | F | HET | POS (HOM) | CC10 x CC18 | 05/10/98 | Cull             |          |     |



|     |   |                |           |             |          |             |          |     |
|-----|---|----------------|-----------|-------------|----------|-------------|----------|-----|
| 169 | F | HET            | POS (HOM) | CC10 x CC18 | 05/10/98 | Cull        |          |     |
| 170 | F | WT             | NEG       | CC10 x CC18 | 05/10/98 | Cull        |          |     |
| 171 | M | HET            | POS (HOM) | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 172 | F | sample missing |           | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 173 | F | sample missing |           | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 174 | F | sample missing |           | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 175 | F | HET            | POS (HOM) | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 176 | M | sample missing |           | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 177 | F | sample missing |           | CC54 x CC60 | 25/10/98 | Cull        |          |     |
| 178 | F | HET            | POS (HOM) | CC10 x CC18 | 01/11/98 | Cull        |          |     |
| 179 | M | HET            | NEG       | CC10 x CC18 | 01/11/98 | Cull        | 12/02/99 | 101 |
| 180 | F | HET            | POS (HOM) | CC10 x CC18 | 01/11/98 | Cull        |          |     |
| 181 | F | HET            | POS (HOM) | CC59 x CC67 | 04/11/98 | Cull        |          |     |
| 182 | M | HET            | POS (HOM) | CC59 x CC67 | 04/11/98 | Cull        |          |     |
| 183 | F | WT             | POS (HOM) | CC59 x CC67 | 04/11/98 | Cull        |          |     |
| 184 | F | HET            | NEG       | CC59 x CC67 | 04/11/98 | Cull        | 12/02/99 | 98  |
| 185 | M | HET            | POS (HOM) | CC59 x CC67 | 04/11/98 | Cull        |          |     |
| 186 | F | HET            | NEG       | CC59 x CC67 | 04/11/98 | Cull        | 12/02/99 | 98  |
| 187 | M | HET            | NEG       | CC55 x CC62 | 07/10/98 | Cull        | 12/02/99 | 125 |
| 188 | F | HET            | POS (HOM) | CC55 x CC62 | 07/10/98 | Cull        |          |     |
| 189 | F | HET            | NEG       | CC55 x CC62 | 07/10/98 | Cull        | 12/02/99 | 125 |
| 190 | F | HET            | POS (HOM) | CC55 x CC62 | 07/10/98 | Cull        |          |     |
| 191 | M | HET            | POS (HOM) | CC55 x CC62 | 07/10/98 | Cull        |          |     |
| 192 | M | HET            | NEG       | CC55 x CC62 | 07/10/98 | Cull        | 12/02/99 | 125 |
| 193 | M | WT             | NEG       | CC54 x CC60 | 23/11/98 | Cull        | 12/02/99 | 79  |
| 194 | M | WT             | NEG       | CC54 x CC60 | 23/11/98 | Cull        | 12/02/99 | 79  |
| 195 | F | WT             | NEG       | CC54 x CC60 | 23/11/98 | Cull        | 12/02/99 | 79  |
| 196 | M | WT             | NEG       | CC54 x CC60 | 23/11/98 | Cull        | 12/02/99 | 79  |
| 197 | M | WT             | NEG       | CC54 x CC60 | 23/11/98 | Cull        | 12/02/99 | 79  |
| 198 | M | WT             | NEG       | CC55 x CC62 | 02/12/98 | Cull        | 12/02/99 | 70  |
| 199 | M | HET            | NEG       | CC55 x CC62 | 02/12/98 | Cull        | 12/02/99 | 70  |
| 200 | M | WT             | POS (HOM) | CC55 x CC62 | 02/12/98 | Cull        |          |     |
| 201 | F | WT             | NEG       | CC55 x CC62 | 02/12/98 | Cull        | 12/02/99 | 70  |
| 202 | F | HET            | POS (HOM) | CC55 x CC62 | 02/12/98 | Gut tumours | 02/08/99 | 240 |
| 203 | F | HET            | POS (HOM) | CC55 x CC62 | 02/12/98 | Gut tumours | 14/06/99 | 192 |
| 204 | M | WT             | POS (HOM) | CC55 x CC62 | 02/12/98 | Cull        |          |     |
| 205 | M | HET            | POS (HOM) | CC55 x CC62 | 02/12/98 |             |          |     |
| 206 | F | HET            | POS (HOM) | CC55 x CC62 | 02/12/98 | Gut tumours | 30/04/99 | 148 |
| 207 | M | sample missing |           | CC10 x CC18 | 26/11/98 | Cull        |          |     |
| 208 | F | HET            | POS (HOM) | CC54 x CC60 | 19/12/98 | Gut tumours | 13/12/99 | 354 |
| 209 | F | HET            | POS (HOM) | CC54 x CC60 | 19/12/98 | Gut tumours | 13/09/99 | 264 |

|     |   |                |           |                 |          |             |          |     |
|-----|---|----------------|-----------|-----------------|----------|-------------|----------|-----|
| 210 | M | HET            | POS (HOM) | CC54 x CC60     | 19/12/98 | Gut tumours | 22/02/00 | 423 |
| 211 | F | HET            | NEG       | CC54 x CC60     | 19/12/98 | Cull        |          |     |
| 212 | F | WT             | NEG       | CC54 x CC60     | 19/12/98 |             |          |     |
| 213 | F | WT             | NEG       | CC54 x CC60     | 19/12/98 |             |          |     |
| 214 | M | HET            | POS (HOM) | CC10 x CC18     | 17/12/98 | Cull        |          |     |
| 215 | F | WT             | POS (HOM) | CC10 x CC18     | 17/12/98 | 1/02/99     | 01/02/99 | 44  |
| 216 | M | HET            | POS (HOM) | CC10 x CC18     | 17/12/98 | Cull        |          |     |
| 217 | M | HET            | POS (HOM) | CC10 x CC18     | 17/12/98 | Cull        |          |     |
| 218 | M | HET            | POS (HOM) | CC55 x CC62     | 26/12/98 | Cull        |          |     |
| 219 | F | HET            | POS (HOM) | CC55 x CC62     | 26/12/98 | Cull        |          |     |
| 220 | F | WT             | POS (HOM) | CC55 x CC62     | 26/12/98 | Gut tumours | 14/06/99 | 168 |
| 221 | F | HET            | POS (HOM) | CC55 x CC62     | 26/12/98 | Cull        |          |     |
| 222 | F | HET            | POS (HOM) | CC55 x CC62     | 26/12/98 | Gut tumours | 18/06/99 | 172 |
| 223 | M | HET            | POS (HOM) | CC55 x CC62     | 26/12/98 | Cull        |          |     |
| 224 | M | WT             | NEG       | CC143 x Min     | 10/01/99 | Cull        |          |     |
| 225 | M | HET            | POS (HOM) | CC143 x Min     | 10/01/99 | Cull        |          |     |
| 226 | M | WT             | POS (HOM) | CC143 x Min     | 10/01/99 | Cull        |          |     |
| 227 | M | HET            | POS (HOM) | CC143 x Min     | 10/01/99 | Breed       |          |     |
| 228 | M | HET            | NEG       | CC143 x Min     | 10/01/99 | Cull        |          |     |
| 229 | F | HET            | NEG       | CC143 x Min     | 10/01/99 |             |          |     |
| 230 | M | WT             | NEG       | CC143 x Min     | 10/01/99 |             |          |     |
| 231 | F | HET            | POS (HOM) | CC143 x Min     | 10/01/99 |             |          |     |
| 232 | M | WT             | POS (HOM) | CC143 x Min     | 04/02/99 | Found dead  | 24/04/00 | 440 |
| 233 | M | HET            | NEG       | CC143 x Min     | 04/02/99 | Gut tumours | 25/04/00 | 165 |
| 234 | M | WT             | NEG       | CC143 x Min     | 04/02/99 |             |          |     |
| 235 | F | HET            | POS (HOM) | CC143 x Min     | 04/02/99 | Gut tumours | 11/05/99 | 97  |
| 236 | F | HET            | NEG       | CC143 x Min     | 04/02/99 | Gut tumours | 12/06/00 | 488 |
| 237 | F | WT             | NEG       | CC143 x Min     | 04/02/99 |             |          |     |
| 238 | M | WT             | POS (HOM) | CC143 x Min     | 04/02/99 | Gut tumours | 27/01/00 | 353 |
| 239 | M | HET            | POS (HOM) | CC143 x Min     | 04/02/99 | Found dead  | 24/04/00 | 440 |
| 240 | M | HET            | POS (HOM) | CC143 x Min     | 04/02/99 |             |          |     |
| 241 | F | HET            | POS (HOM) | CC143 x Min     | 20/04/99 | Gut tumours | 16/11/99 | 206 |
| 242 | F | WT             | POS (HOM) | CC143 x Min     | 20/04/99 |             |          |     |
| 243 | F | HET            | POS (HOM) | CC143 x Min     | 20/04/99 | Gut tumours | 28/01/00 | 278 |
| 244 | M | HET            | NEG       | CC143 x Min     | 20/04/99 | Cull        |          |     |
| 245 | M | HET            | POS (HOM) | CC143 x Min     | 20/04/99 |             |          |     |
| 246 | F | HET            | NEG       | CC143 x Min     | 20/04/99 | Gut tumours | 19/10/99 | 179 |
| 247 | M | HET            | POS (HOM) | CC227 x 231,241 | 07/06/99 | Gut tumours | 30/11/99 | 173 |
| 248 | M | HET            | NEG       | CC227 x 231,241 | 07/06/99 | Found dead  | 17/01/00 | 220 |
| 249 | F | WT             | POS (HOM) | CC227 x 231,241 | 07/06/99 |             |          |     |
| 250 | M | WT             | POS (HOM) | CC227 x 231,241 | 07/06/99 |             |          |     |
| 251 | F | sample missing |           | CC227 x 231,241 | 12/06/99 |             |          |     |
| 252 | F | WT             | NEG       | CC227 x 231,241 | 12/06/99 |             |          |     |



|     |   |     |           |                 |          |             |          |     |
|-----|---|-----|-----------|-----------------|----------|-------------|----------|-----|
| 253 | M | HET | POS (HOM) | CC227 x 231,241 | 12/06/99 |             |          |     |
| 254 | M | HET | NEG       | CC227 x 231,241 | 27/07/99 |             |          |     |
| 255 | F | HET | NEG       | CC227 x 231,241 | 27/07/99 | Cull        |          |     |
| 256 | F | WT  | POS (HOM) | CC227 x 231,241 | 27/07/99 | Cull        |          |     |
| 257 | F | HET | POS (HOM) | CC227 x 231,241 | 27/07/99 |             |          |     |
| 258 | M | HET | NEG       | CC227 x 231,241 | 27/07/99 |             |          |     |
| 259 | F | WT  | POS (HOM) | CC227 x 231,241 | 07/08/99 | Found dead  | 24/04/00 | 267 |
| 260 | F | HET | POS (HOM) | CC227 x 231,241 | 07/08/99 | Gut tumours | 12/06/00 | 315 |
| 261 | F | HET | POS (HOM) | CC227 x 231,241 | 07/08/99 | Cull        |          |     |
| 262 | F | HET | POS (HOM) | CC227 x 231,241 | 07/08/99 | Gut tumours | 12/06/00 | 305 |
| 263 | M | HET | POS (HOM) | CC227 x 231,241 | 07/08/99 |             |          |     |
| 264 | M | WT  | POS (HOM) | CC227 x 231,241 | 07/08/99 |             |          |     |
| 265 | M | HET | NEG       | CC227 x CC231   | 22/09/99 |             |          |     |
| 266 | F | HET | POS (HOM) | CC227 x CC231   | 22/09/99 |             |          |     |
| 267 | F | HET | NEG       | CC227 x CC231   | 22/09/99 |             |          |     |
| 268 | F | HET | NEG       | CC227 x CC231   | 22/09/99 |             |          |     |
| 269 | F | HET | NEG       | CC227 x CC231   | 22/09/99 |             |          |     |
| 270 | M | WT  | NEG       | CC227 x CC231   | 22/09/99 |             |          |     |
| 271 | F | HET | POS (HET) | Min x CC91,CC61 | 22/09/99 |             |          |     |
| 272 | M | WT  | POS (HET) | Min x CC91,CC61 | 16/10/99 |             |          |     |
| 273 | M | WT  | NEG       | Min x CC91,CC61 | 16/10/99 |             |          |     |
| 274 | M | WT  | NEG       | Min x CC91,CC61 | 16/10/99 |             |          |     |
| 275 | M | HET | POS (HOM) | CC227 x CC231   | 06/11/99 |             |          |     |
| 276 | M | WT  | POS (HOM) | CC227 x CC231   | 06/11/99 |             |          |     |
| 277 | F | HET | POS (HOM) | CC227 x CC231   | 06/11/99 | Gut tumours | 27/11/00 | 381 |
| 278 | F | HET | POS (HOM) | CC227 x CC231   | 06/11/99 | Gut tumours | 17/05/00 | 191 |
| 279 | F | HET | POS (HOM) | CC227 x CC231   | 06/11/99 |             |          |     |
| 280 | F | HET | NEG       | CC227 x CC231   | 01/12/99 | Found dead  | 10/07/00 | 219 |
| 281 | M | WT  | POS (HOM) | CC227 x CC231   | 01/12/99 |             |          |     |
| 282 | F | HET | NEG       | CC227 x CC231   | 01/12/99 |             |          |     |
| 283 | M | WT  | NEG       | CC250 x CC256   | 18/03/00 |             |          |     |
| 284 | M | WT  | NEG       | CC250 x CC256   | 18/03/00 |             |          |     |
| 285 | F | WT  | NEG       | CC250 x CC256   | 18/03/00 |             |          |     |
| 286 | F | WT  | NEG       | CC250 x CC256   | 18/03/00 |             |          |     |

#### CC mouse database legend

| Apc status          | Transgene status  |
|---------------------|---|
| HET = $Apc^{Min/+}$ | NEG = Transgene negative by PCR                                       |
| WT = $Apc^{+/+}$    | POS (HET) = Transgene positive by PCR, one transgene positive parent  |
|                     | POS (HOM) = Transgene positive by PCR, two transgene positive parents |